

Kåre Nitter Rugesæter

*Associate professor, Faculty of Education
Bergen University College*

Difficult Contrasts: an analysis of phonemic distinctions in the English of young Norwegian learners seen against the backdrop of incidental foreign-language learning

Abstract

Norwegian children in the 21st century are quite massively exposed to English as an L2, primarily through subtitled television programmes, films and various computer-based activities. The article investigates to what extent this passive exposure influences their basic phonological competence in the L2 by analyzing their ability to actively distinguish between certain phoneme pairs in English which are considered difficult for Norwegian learners, mainly due to interference from the phonemic system in their L1. The data material was collected through 136 recordings of 11-16 year old pupils in ten Norwegian state schools. By comparing the results to similar studies made in Norway 20-35 years ago, at a time when young people's exposure to English was considerably more limited, the article discusses the effects of incidental foreign language acquisition on young people's L2 competence. The discussion is also placed in the light of recent, international research in this field. By critically investigating some of the findings of that research, the article furthermore attempts at problematizing the concept of L2 acquisition by focusing on the distinction between active and passive language competence. The findings indicate that basic L2 phonological competence is in no way automatically enhanced by passive exposure to the language. The article points to the necessity of structured teaching programmes to counterweight the complexity and variations in the L2 input that young people encounter in their media based exposure to the L2.

Keywords: *English phonetics and phonology; incidental language learning; teaching English as an L2*

Introduction

Children and young people in Norway today are quite massively exposed to the English language, particularly through the media, and clearly more so than earlier generations. This article investigates to what extent this exposure has

implications for aspects of young learners' basic oral acquisition. The focus is primarily on phonological contrasts which have been considered difficult for the learners due to first language (L1) transfer, but which will be present in the sound inventory of native speakers (NS) of English. The article includes a discussion on incidental foreign-language acquisition through media exposure, with reference to passive versus active language skills and foreign-language acquisition in general, by comparing its findings to similar, earlier studies in Norway (Handeland, 1987; Hestetræet, 1993; Johannessen, 1979; Nilsen, 1989; Rugesæter, 1992), as well as relevant studies carried out elsewhere in Europe (d'Ydewalle & Van de Poel, 1999; Kuppens, 2010; Lefever, 2012; Leppink, 2010; Mitterer & McQueen, 2009; Vanderplank, 2009). The article also touches upon how its findings may affect the methods and foci for the teaching of English as an L2 to young learners.

Four phonological contrasts were investigated in this study: between /s/ and /z/, between /ɪə/ and /eə/ (US /ɪr/ and /er/), between /e/ and /æ/ and between /əʊ/ (US /oo/) and /aʊ/. The main reason for choosing these contrasts is that they have been investigated in earlier studies of Norwegian learners' phonemic inventory, something which facilitates a comparative, longitudinal approach. The findings of the present study may be divided into two main categories: sound contrasts that clearly create problems for young learners, notably the /s-z/ and the /ɪə-eə/ (/ɪr-er/) contrasts; and those that seem to create relatively few problems, i.e. the /e-æ/ and, in particular, the /əʊ/ (/oo/)-aʊ/ contrasts. It is interesting to note that all these findings seem to correspond well with the findings of earlier studies (Handeland, 1987; Hestetræet, 1993; Johannessen, 1979; Nilsen, 1989; Rugesæter, 1992), indicating that young learners' phonological acquisition does not seem to have changed much, in spite of the substantial increase in passive exposure to the L2. The findings do not, however, support the conclusions in a number of studies on foreign-language learning based on measuring the effects of passive input through media exposure (Bird & Williams, 2002; Koolstra & Beentjes, 1999; Kuppens, 2010; Lefever, Ragnarsdóttir, & Torfadóttir, 2006; Leppink, 2010; Markham, 1999; Torfadóttir, 1994; Williams and Thorne, 2000). These studies have investigated various aspects of the development of language competence in L2 learners, such as their understanding of written and spoken language, oral production related to vocabulary building, and to a certain extent grammar skills and conversational skills. However, they have only dealt with phonological acquisition, which is the focus of this study, to a very limited extent, if at all.

Language learning through the type of passive input that one is exposed to through the media (and other types of passive input not followed up by practicing the language) has been labelled *incidental foreign-language acquisition* (see, for example, d'Ydewalle & Van de Poel, 1999; Vanderplank, 2009). It has been assumed that such input will help in building competence in a foreign language, and this assumption has been supported by research.

According to Kuppens (2010), several empirical studies have found that children and adolescents can incidentally learn a foreign language through watching television. In particular, research on incidental acquisition through watching subtitled television programmes has been carried out in countries with relatively small language communities - for example Holland, Flemish-speaking Belgium and Iceland - where subtitling rather than dubbing is common, largely due to economic considerations related to the size of the target group (for an overview of this research, see, for example, Lefever, 2012 and Vanderplank, 2009, pp.12-22).

The general use and effects of modern audio-visual and digital technology in language teaching, as well as teachers' attitudes towards this new technology, has been the subject of considerable critical research internationally for a number of years; see, for example, Bush, 2000; Grgurovic & Hegelheimer, 2007; Hulstijn, 2003; Koolstra & Beentjes, 1999; Lam, 2000 and Tschirner, 2001. Kramsch and Anderson's (1999) study of the effects of the use of audio-visual, multi-media technology to reflect target-language culture as well as linguistic input is an example of the generally optimistic tone in this research regarding the potential of new technology in language learning. Similarly, Weyers's (1999) article on television programmes as a valuable asset in providing L2 input as part of a structured teaching programme highlights the positive effects of such methods. Herron, York, Corrie and Cole's (2006) study comparing the learning outcomes of two groups of students, one taking part in a film-based course and the other participating in a comparable text-based course, showed that the former group had a better learning outcome overall.

Studies which have focussed on the effects of visual and auditory input with and without captions and subtitles have generally suggested that learners benefit from the combination of text, subtitles and visual information; see, for example, Bird & Williams, 2002; Huang and Eskey, 1999-2000; Markham, 1989 and Williams & Thorne, 2000. However, the conclusions from some more recent studies have questioned some of these effects (Taylor, 2005) or shown limited improvement among the learners (Yuksel & Tanriverdi, 2009).

Other recent studies have investigated the role of media input with regard to L1 and language change (Stuart-Smith, 2011), stressing, among other things, the high degree of variation in the language expressions found in the media today. This type of research contributes to the understanding of the complexity of incidental language learning through watching television, aspects of which will be discussed later in this article in relation to the findings of the present study.

The complex discussion related to second-language acquisition and age has largely centred on the *sensitive language hypothesis theory* (see, for example, Doughty & Long, 2003, pp. 539-588 and Lenneberg, 1967). However, there seems to be general agreement that from approximately two years of age and up to puberty children attain L2 skills more easily than older learners. This will include, not least, basic, phonological acquisition. Adults acquiring a second

language tend, for example, to retain a non-native accent in that language (Archibald, 1998, p. 37). This is confirmed by central research investigating phonological acquisition in particular: “Age of L2 learning appears to be the most important predictor of degree of foreign accent.” (Piske, MacKay, & Flege, 2001, p. 191). The vast majority of the informants in the present study were between eleven and thirteen years old at the time of the recordings, i.e. within an age group that is particularly sensitive to language learning. The findings in the present study, however, indicate that passive exposure to an L2 through rather heavy media exposure over time, where subtitled television programmes play a central role, has minimal, or at least very limited, effect on phonological acquisition among young learners.

Method

Experimental design

With the aim of investigating whether Norwegian children today are closer to L1 phonological patterns in English than earlier generations, 136 eleven to sixteen-year-old pupils in ten schools in western Norway were asked to read a simple text and a set of specially designed sentences in English, and to answer a set of questions about their overall exposure to English. Emphasis was put on features where Norwegian learners are known to have special problems, either due to transfer from their L1 sound system or because of speech sounds and contrasts that do not exist in Norwegian (Nilsen, 2002; Nilsen & Rugesæther, 2008). The focus was not on whether their pronunciation as such is close to a specific NS norm, but whether contrasts that are present in all major L1 accents are present in the pronunciation profile of these children. It is taken for granted that children have adopted a full set of phonological contrasts in their first language by the age of eleven. The study comprises eight English speech sounds (approximately 20% of the sound inventory in English). Using Wells’s (1982) classification system, the main focus has been on what he labels *systemic* differences.

The features investigated are the contrasts between the centring diphthongs /ɪə/ as in <beer> and /eə/ as in <bear> (represented as a contrast between the combinations /ɪr/ and /er/ respectively in rhotic accents); the monophthongs /e/ as in <bed> and /æ/ as in <bad>; and the diphthongs /əʊ/ (US /oʊ/) as in <no> and /aʊ/ as in <now>. In addition, the study investigates the contrast between /s/ and /z/, as it comes out in word pairs such as <ice> and <eyes> or <Sue> and <zoo>; as well as the pronunciation of /z/ between voiced segments, as in <present>.

In a short, structured interview (in Norwegian, before the recording started), each informant was asked several questions about his or her media habits. The first was ‘How often do you watch television programmes where English is used?’, and they were presented with the following alternatives:

- Every day or nearly every day
- Four or five days a week
- Two or three days a week
- Once a week or less

In addition, they were all asked several follow-up questions:

- Can you mention some of the programmes that you watch?
- Do you watch English films on DVD?
- Do you hear English through computer games and / or the internet?
- Are you exposed to English in other ways (personal relationships / holiday trips, etc.)?
- How often do you speak English yourself (at school and in other contexts)?

The informants

At the time of the recording, the youngest informants (50 in all) were in the fifth grade in the Norwegian state school system; the second group (74 in all) were seventh graders; and the oldest pupils (12 in all) were in the tenth grade. All these students had started to learn English at school in the first grade, when they were six. The recordings took place in 2010-2012 using a small digital recorder at the schools the informants normally attended, i.e. in an environment that was familiar to them, in order to reduce the formality of the situation as much as possible. The informants were recruited on a voluntary basis through their form teacher. A clear majority of the pupils in all of the classes participated, but pupils with special linguistic backgrounds were excluded, for example immigrant children with a first language other than Norwegian, and pupils with reading disabilities which made it difficult for them to take part.

Since 1997, English has been a compulsory subject in Norwegian state schools from grade one, i.e. pupils start their formal education in English from the age of six. During the first two years, the number of lessons is limited; according to the national norm (*Læreplanverket for Kunnslopsløftet*, 2006) they have a total of 138 lessons in the course of the first four years and 228 lessons during the next three, years five to seven. The majority of the informants therefore had had between 150 and 350 English lessons at the time of the recording.

Basis for comparison

In order to investigate whether there has been a development over the last few decades with regard to Norwegian learners' phonological acquisition in English, it is necessary to compare the results of the present findings with earlier studies. There are a few studies that may be used as a basis for comparison. Only two of these (Hestetraet, 1993 and Rugesæter, 1992) actually focus on phonological contrasts; however, there are others (Handeland, 1987; Johannessen, 1979;

Nilsen, 1989) dealing with phonetic descriptions that can be used indirectly by interpreting the results with phonological contrasts in mind. Handeland (1987) and Nilsen (1989) limit their analyses to vowel sounds only, whereas the other analyses cover all the phonemes dealt with in this article. Although the focus and method of these studies varied somewhat, the data material in all of them was based on the informants reading a text in English. The studies centred on learners who were generally in their late teens or (most of them) early twenties at the time of the recordings. The overall exposure of young people to English when these studies were conducted was much more limited than is the case with children and young people today, due to the introduction of cable and satellite television from the 1990s onwards, and of the internet services that are now a common feature in practically every home in Norway.

Auditory analysis

The descriptions below are based on auditory analysis. The main focus is on contrasts, and the reading text therefore contained a number of minimal or near-minimal pairs to avoid allophonic variation as much as possible. To avoid making the contrasts stand out too much in the text structure, they were regularly made less visible by not putting them side by side in isolation. The idea was that by placing the relevant words in a context, it would be easier for the informants to identify their meaning, particularly since the rather unpredictable system of spelling representation in English does not always make it clear, not least to a foreign learner, what the words mean when seen in isolation.

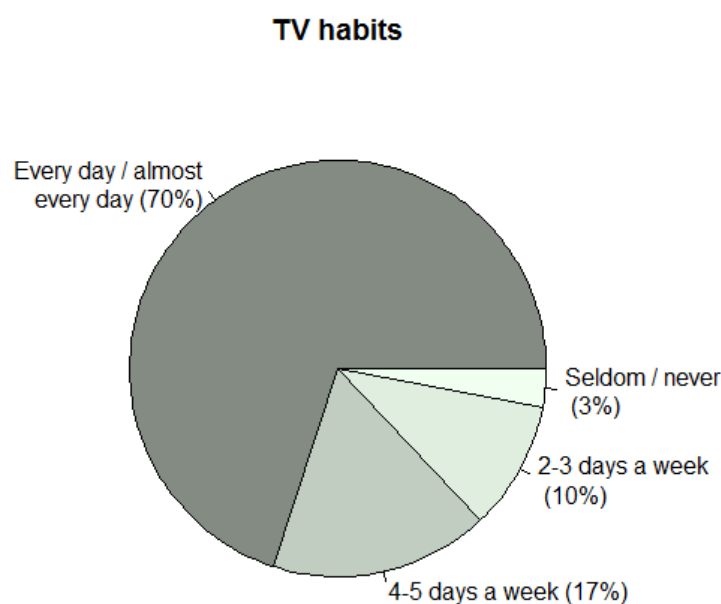
Auditory analysis has its potential weaknesses. It is not always easy to hear whether speech sounds are the same or not. In actual fact, two individual sounds are phonetically never 100% identical. However, it is normally possible to hear whether the given phonemic contrasts are present or not with native speakers and non-native speakers of a language alike. There are bound to be borderline cases and grey areas, not least because of allophonic variations, and ambiguous or unclear cases have been marked as such in the data. On the whole, however, in the framework of the present study, the kind of auditory analysis used for this article must be considered an acceptable tool, and a combination of features (vowel duration as well as quality, voicing as well as lenition in various phonetic contexts) has been used to establish whether a potential phonemic contrast exists, or not. With as many as 136 informants, it is possible to identify certain rather clear-cut indicators. Furthermore, it may be argued that distinctive contrasts that cannot be perceived in an auditory analysis are linguistically irrelevant.

In an ideal world, it would have been desirable to compare the findings of the present study to previous studies with informants of the same age group and with informants with the same L1 variety. It would also have been desirable to compare the findings to studies that made use of exactly the same methodology. However, such studies are not available, and in view of the dominant theories on

language learning and age, such as the sensitive language hypothesis theory (Lenneberg, 1967), a comparison of 11-16 year-old language learners with learners in their late teens and early twenties should still provide us with relevant information about changes regarding phonological acquisition related to increased exposure to an L2. A limitation in this study lies in the fact that only one specific, albeit fundamental, aspect of oral L2 acquisition is measured.

Findings

Figure 1: Overview over how much the informants watched TV. N=136.



The interviews: Exposure to English through the media

As seen in Figure 1, 70% of the 136 children said that they watched television with programmes in English every day or almost every day; 17% stated that they watched such programmes 4-5 days a week; and 10% said that they watched such programmes 2-3 days a week. Only 3% maintained that they watched television with programmes in English once a week or less. Even though the number of hours these children spend in front of the television or computer screen may vary, and the information they provided in this kind of interview situation may not be 100% accurate, the figures below clearly function as an indicator of their passive exposure to English through modern media, primarily television. Some English-language television programmes for children in Norway are dubbed, but when the informants were asked what sort of programmes they watched, the most typical answers in this study were sitcoms, soaps and other series mainly targeted at a teenage or adult audience, plus Animal Planet, and for many boys, Discovery Channel. None of the programmes

mentioned are dubbed when transmitted in Norway; all of them are merely subtitled.

One might conclude that almost 90% of the informants had been exposed to English on a more or less daily basis through television alone.

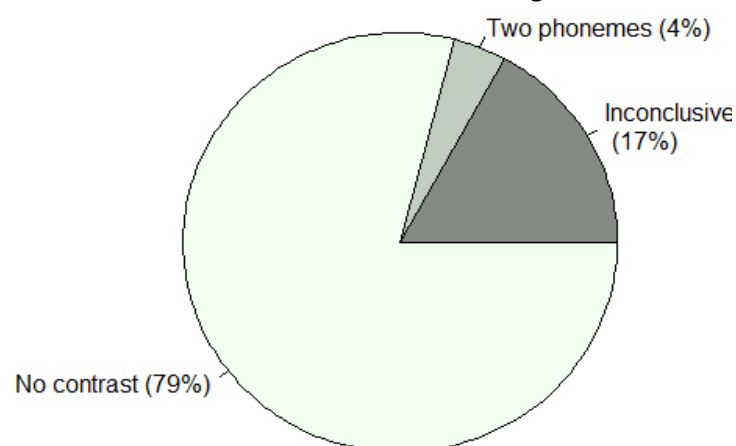
In addition, these children would to a certain, although varying, degree have been exposed to English through other media, not least through the use of computers. Admittedly, much of this would be English in its written form, but there is a growing use of multimedia functions not only in computer games, but also through sound files and video clips on the web. When asked about sources other than television where they encounter English, more than 80% of the children mentioned at least one additional source - e.g. films / DVDs, computer games with sound, music, or holiday visits to English-speaking countries - all of which add to young Norwegians' overall exposure to oral English these days.

There is every reason to believe that 11-16 year olds today are exposed to more spoken English than children 30-50 years ago, when the media situation was very different. In those days there was, for example, only one TV channel in Norway, and most of the programmes were in Norwegian. No households had access to English-speaking TV channels via satellite and cable. The type of multimedia resources that have become a natural part of young people's everyday life today, including the Internet, were unknown to children in those days. In addition, at that time they did not start to learn English at school until the age of eleven (or later), compared to six today.

Phonemic analyses

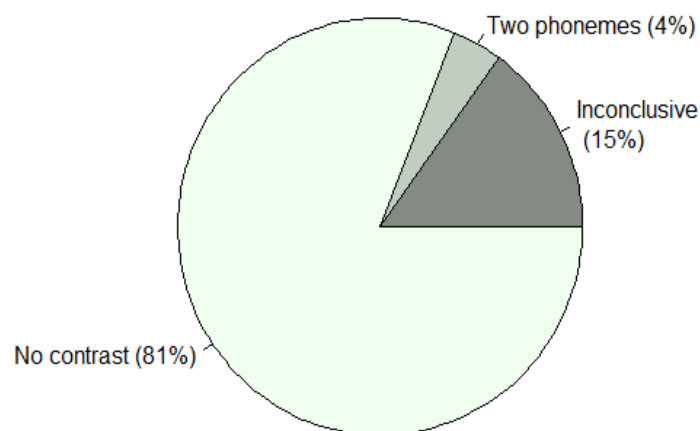
In the following, the auditory analyses of the informants' phonemic contrasts, or lack of such, are presented.

Figure 2: Overview over whether the informants distinguished between /s/ and z/



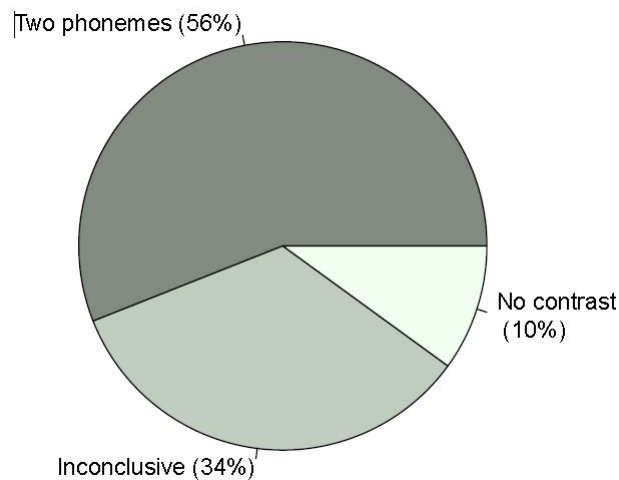
As seen in Figure 2, less than 5% of my informants made a clear and apparently systematic distinction between /s/ and /z/ in their reading. Approximately 80% showed no signs of distinguishing between the two at all. That means that they used a voiceless fricative, /s/, throughout, including words where the spelling was <z>, as exemplified by the word <zoo>. <ice> and <eyes> are homophonous and the /z/ in a word like <present> shows no sign of lenition or voicing. A couple of informants pronounced <zoo> with a voiced fricative, reflecting an awareness of the contrast when the spelling is <z>, but not in other cases. Somewhat less than 20% of the informants showed signs of vowel lengthening in the words <lies> and <eyes>, which *may* be due to influence from the <e> in spelling following the first vowel letter, but since there are no other signs of a distinction, I have put those down as inconclusive. Nevertheless, the overall impression for the group as a whole is clear: the vast majority have not acquired the contrast in question, and replace the /z/ with the nearest L1 equivalent (Norwegian does not have /z/ or any other lenis fricatives.)

Figure 3: Overview over whether the informants distinguished between /ɪə/-/eə/ (/ɪr /- /er/)



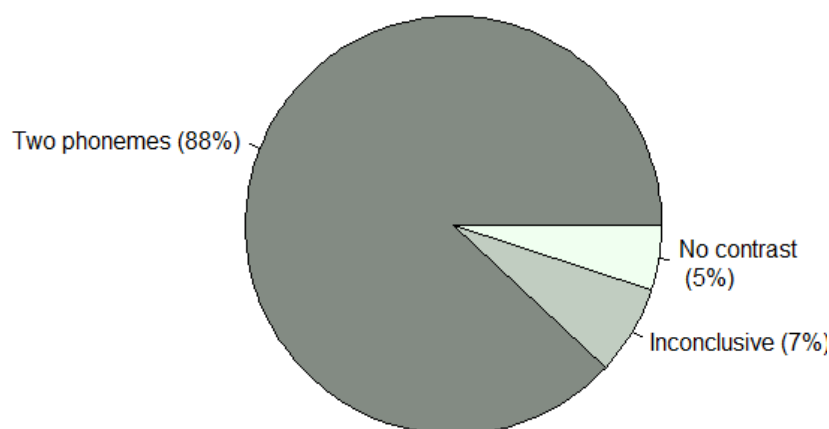
As shown in Figure 3, a little over 80% of my informants showed no signs of distinguishing between these two phonemes, so that <ear> and <air> were homophonous. For approximately 15% of them, the findings were inconclusive. In some cases there would probably be some confusion among the informants due to interference from overlapping spelling, for example where <ea> is used for both phonemes; this may complicate the acquisition of this phonemic contrast, but only a little over 4% made a clear distinction between the two phonemes.

Figure 4: Overview over whether the informants distinguished between /e/ and /æ/



As Figure 4 shows, more than half (56%) of my informants made a clear and consistent distinction between these two sounds. With a quarter of this group, the distinction seemed to be based solely, or at least primarily, on increased length when the phonemes occurred before lenis consonants, as in <bed> and <bad>, with little or no audible difference in quality. The existence / non-existence of a contrast among a third of the group (34%) had to be classified as inconclusive, either because any difference was difficult to hear, or because of a lack of consistency that made it difficult to know whether they made a systematic distinction between the two sounds. 10% clearly did *not* make any distinction between the two phonemes in their speech. It should be added that this contrast differs from the other phoneme pairs in this study in that it is recognizable from a similar, or more or less parallel, contrast in the informants' L1.

Figure 5: Overview over whether the informants distinguished between /əʊ/ (/ou/) and /aʊ/



In spite of all the potential problems related to phonetic closeness, overlapping spelling representations and L1 variations concerning these two diphthongs, this contrast does not seem to create any difficulties for young learners today, as can be seen in Figure 5. Just under 5% failed to make any distinction between the two, and 7 % fell into the category “inconclusive”, since it was difficult to properly judge whether both phonemes were represented. That means that 88% of the informants made a clear distinction between the two phonemes.

Comparison to earlier studies

The findings regarding the four phonological contrasts were compared to five previous studies carried out in Norway in the late 70s, 80s and early 90s (Handeland, 1987; Hestetraet, 1993; Johannessen, 1979; Nilsen, 1989; Rugesæter, 1992) in order to identify any development or changes in the learners' patterns regarding these L1 distinctions.

/s - z/:

Johannessen's (1979) phonetic descriptions of the pronunciation of these two phonemes among his informants indicate that quite a few, but not the majority, of the students used /s/ for /z/, implying that more than half operated with two phonemes. Nilsen's (1989) statistics show that 85% of his students pronounced /z/ as /s/, thus indicating that the vast majority of them did not distinguish between fortis /s/ and lenis /z/. Likewise, Rugesæter (1992) reports that 65% of his informants twenty years ago made no distinction at all between the two phonemes.

In other words, there seems to be a fairly straightforward correlation between these earlier findings and the findings of the current study, with the exception of Johannessen's study, which indicates better skills on this point among his informants. The lack of any contrast between /s/ and /z/, or rather the lack of a /z/-phoneme, continues to be part of the phonological pattern of young Norwegian learners, although they have been heavily exposed to English during their early years. However, it should be added that since some of the informants in the present study had adopted a /z/ where the spelling was <z>, they may have showed possible signs of a process where the /z/ was being incorporated into their phonemic inventory on a more general basis.

/ɪə - eə/ - (/ɪr - er/)

From the available data, it is possible to conclude that a great many of the informants in all the previous studies had problems with this contrast. Handeland (1987, p. 32), reports that only 19% of her informants distinguished between these two speech sounds. However, she also found that 176 students (35%) pronounced /ɪə / as /eə/, which suggests that the number of students who

lacked a contrast here might actually have been higher; in a comment, she states that the most common mistake when it came to these two phonemes was to use [ɛ] as a the “first element for both” (Handeland, 1987, p. 34). Nilsen’s (1989) study suggests that about two thirds of his informants failed to make a distinction between these two speech sounds. Rugesæter (1992) reports similar findings, with 65% of his students not distinguishing between these vowel phonemes; in addition, approximately 25% had pronunciation weaknesses that made it difficult to assess whether they actually had this contrast systematically incorporated into their profile, and less than 10% made a clear, systematic distinction between them. Hestetraet’s (1993) findings further support this impression. She has a more careful approach in her analysis, which means that she is reluctant to rule out any possible phonemic distinction based on her material. She has not found one single informant who clearly and systematically operated with two phonemes, and claims that 95% either clearly made no distinction between the two, or that it was “doubtful” that they did. There is no indication in the findings of the present study that young people today are much better at actively distinguishing this sound pair in English compared to Norwegian learners two or three decades ago.

/e/ - /æ/:

Johannessen’s (1979) study is based on a limited number of informants (12) and is inconclusive when it comes to this sound pair. Like Johannessen, Handeland (1987) did not investigate specifically whether her informants had incorporated the contrast between /e/ and /æ/ in their English pronunciation. Judging from her descriptions of sound manifestations, one might assume that only a small percentage (perhaps 4% or possibly a little more) failed to make a distinction between the two phonemes. Nilsen (1989) reports in his study of 172 students that 59% of them had an /e/-phoneme which was ‘too open’ (compared to RP), while 5% had an /æ/-phoneme which was ‘not open enough’. To what extent any of his informants actually failed to distinguish between /e/ and /æ/ at all is unclear, but since a lack of distinction between /əʊ/ and /aʊ/ is reported for some of his informants, one can assume that a similar coalescence between /e/ and /æ/ was not registered, or at least not a focal point in the analysis. Rugesæter (1992) found that about 40% of his informants did *not* make a distinction between the two phonemes, so that <said> and <sad> were homophonous. Hestetraet (1993), however, found that 52% of her 42 informants made a clear distinction between the two vowel phonemes, whereas 45% were characterized as unclear (‘doubtful’ / ‘inconsistent’ / ‘blurred’); only 1 out of 42 clearly did *not* have two phonemes in his/her pronunciation profile.

In spite of slightly varying results, it must be said, looking at the five previous studies as a whole, that 11-13 year olds today seem to have much the same pattern in their pronunciation of English /e/ and /æ/ as was the case with

Norwegian learners of English 20-30 years ago, including a contrast, or lack of such, between these two monophthongs.

/əʊ (/oʊ/) - aʊ/:

Johannessen (1979) concentrates on the phonetic realisations of the two phonemes. He reports that many of his informants had realisations of both phonemes that deviated from the norm (RP), but by studying the phonetic descriptions one can quite safely deduce that no more than 16% of his informants had the same manifestation of both phonemes. Handeland's (1987) findings are relatively difficult to interpret in relation to a potential contrastive loss when it comes to these phonemes, but again by studying the descriptions of the realisations of each phoneme, one may deduce that up to 58 students (11%) failed to make a distinction between the two. Although she reports that some students (an additional 5% with /əʊ/ and 7% with /aʊ/) 'were unstable in their pronunciation' (Handeland, 1987, pp. 30-31), which may be explained as distributional problems, the overall picture seems to be that only a small minority of the students had failed to develop two distinctive phonemes here, in spite of realisational deviations from the norm of reference. Nilsen (1989) reports that 11% of his informants had an /əʊ/-phoneme 'not distinguished from /aʊ/', and 14% had an /aʊ/-phoneme 'not distinguished from /əʊ/'. It is unclear whether these weaknesses apply to the same students, but it seems evident from his data that a clear majority of his informants had two distinct phonemes for /əʊ/ and /aʊ/. Rugesæter (1992) states that more than 80% of his informants had two phonemes for these speech sounds. Hestetraet (1993) reports findings that are somewhat less conclusive. While she does not exclude completely the possibility of any of her informants having two phonemes in this field, she characterizes many of her findings as blurred or doubtful. Her findings do not clash with Rugesæter's (1992, p. 3) conclusion that 'with regard to /əʊ/ and aʊ/ in general, realisational 'inaccuracies' and distributional confusion rather than systemic neutralization seems to be typical with [the informants] as a whole.'

Discussion

The following discussion considers recent research into the effects of incidental language learning through the media, and relates the results of that research to the findings of the present study. The conclusion also considers possible implications of the findings for the teaching of pronunciation in the L2 English classroom and in teacher education.

The type of input a learner receives through watching L1 subtitled television programmes contains an auditory element as well as visual support (pictorial information). The visual support can be divided into two types: the situation in

which any dialogue takes place, and facial expressions and gestures; both may contribute to clarifying and shedding light on linguistic meaning, in addition to the support given by the subtitles themselves via the learners' L1. In theory, this is a favourable context for language learning, and a substantial number of studies indicate the positive learning effects of this type of input. In natural L1 acquisition, however, there is a mixture of input and language practice accompanied by direct and indirect feedback from the learners' surroundings in an interactive process. Exposure to an L2 through the media will not include active use of the language under guidance, only passive reception. It may therefore be relevant to consider the limitations of incidental language learning as well as its positive effects.

When discussing language acquisition, it is useful to distinguish between different types of skills. The basic division into the four language skills (listening, reading, speaking and writing), often referred to explicitly or implicitly in theories and methodology related to language teaching and learning, (e.g. Cook, 2013; Macaro, 2003; Pinter, 2011 and Simensen, 1998, pp. 139-219) is a good starting point. It is natural to assume that passive exposure (listening) to an L2 first and foremost leads to better receptive skills; i.e. by being exposed to the spoken language learners will develop their understanding of what is being said, and in the case of subtitled television, there is the added element of potentially combining the meaning of words and expressions in the L2 and L1. So in developing listening skills and a passive vocabulary (possibly with a reference to the L1 through subtitling), media exposure may help. When it comes to the development of active language skills (e.g. pronunciation skills and grammatical skills), the effect of media exposure may be more uncertain. Leppink's (2010) study, which aims at finding out whether or not exposure to L2 media input benefits L2 proficiency, compares two groups of high school students, one 'high exposure group' and one 'low exposure group'. The experimental design includes a QPT (quick placement test) and a pronunciation task. Her conclusion is that high media exposure had a significant effect on the QPT results, but 'the difference between the two groups on the pronunciation task was not significant' (Leppink, 2010, p. 40). Potentially, she attributes that to the limitations of the experiment, and the type of pronunciation test used, but it may also be connected to the fact that the QPT measures understanding (receptive skills) and has a multiple choice or a cloze test format, whereas pronunciation is part of the active language production. Leppink's group of 'high exposure' informants belong to the same age group as the informants of the present study, and the results are therefore comparable. Thus, it is interesting that the findings are basically the same. Closer examination of some of the more optimistic articles concerning the effects of the combination of auditory and visual input in language learning (Bird & Williams, 2002; Huang and Eskey, 1999-2000; Markham, 1989 and Williams & Thorne, 2000) reveals that they

also measure receptive skills, notably listening comprehension, and not much regarding acquisition linked to the learners' long-term, productive skills.

The QPTs used in Leppink's (2010) experiment may also be said to measure immediate understanding (short-term effects) where the informants may be helped by visual information there and then; in contrast, pronunciation skills are to a large extent a result of a long-term development, and may be more difficult to improve or change in the short term. This perspective is discussed by Kuppens (2010). In her article she 'investigates the extent to which children's foreign language skills benefit from their long consumption of media' (Kuppens, 2010, p. 65). The informants are Flemish Dutch-speaking children in their final year of primary education, and are thus comparable to the informants of the present study in terms of age. Her findings show clear and positive long-term effects from watching subtitled movies and television when it comes to translation skills both ways (Dutch to English and English to Dutch). The children were tested orally, but not on pronunciation. The extent to which the test actually measured their active language proficiency may be somewhat limited. The words and phrases that constituted the texts the children were asked to translate were (understandably enough, given their age) short and simple and partly translucent in terms of L1 – L2 meaning. However, grammatical correctness was included in the scores, so a certain, although limited, degree of active expression in addition to recognition was included in the experimental design.

Earlier research (see, for example, d'Ydewalle & Van de Poel, 1999; Vanderplank, 2009) has pointed to significant effects from media exposure to an L2 on language acquisition among young learners, although again it may be argued that much of what has been measured is linked to passive understanding rather than active performance. Certain active skills, such as pronunciation, i.e. avoiding L1 transfer on the phonological as well as the phonetic level (Leppink, 2010; Rugesæter, 1995), seem to be less influenced by this exposure, even with young learners in their most receptive years of language acquisition.

The limitations of the effect of media exposure have been investigated in earlier research on incidental foreign-language learning, too. One of the weaknesses of television as a medium for language acquisition is related to the overwhelming amount of information, not limited to pure linguistic information, constantly thrown at the recipients. '[...] the language of broadcast programmes comes too quickly for learners and there is too much of it. It is all too often culturally bound or bound in contexts that may be unfamiliar to foreign language learners' (Vanderplank, 2009, p. 9).

In addition, when children watch television in their spare time, they are not in a 'learning mode'; they associate television with leisure activities and are thus not receptive to language input to the same extent as in a classroom situation. Weyers's (1999) study shows that television *is* an asset in foreign language learning, when used as part of a structured class activity.

With regard to subtitled television programmes in particular, there is evidence that young learners do not benefit much from their age-related receptive skills compared to adult learners in relation to the overall learning effect. d'Ydewalle & Van de Poel (1999), in their assessment of the learning effect of incidental foreign-language acquisition among adult learners and children (Dutch 8-12 year olds) through watching subtitled television programmes in a foreign language, claim that '[they] obtained real but limited foreign-language acquisition', but that 'the learning of the children was not superior to adults investigated in prior studies.' (p. 227) This is, of course, not in accordance with the sensitive language hypothesis theory. The findings may, to some extent at least, be explained by the limited reading skills of the children, although eye movement research (d'Ydewalle & Van Rensbergen, 1989) shows that children down to the age of eight seem to have the same level of attention as adults when watching subtitled television programmes. Earlier research also suggests that children make an effort to process the sound track (e.g. Pezdek & Hartman, 1983; Pezdek & Stevens, 1984).

A recent study by Mitterer and McQueen (2009) into learning via subtitled television produced some interesting findings that can be directly linked to the findings of the present study. The usual design for subtitled television programmes is an L2 (the original) soundtrack with L1 (native-language) subtitles. Another combination, found mostly on DVDs /Blue-Ray movies, is an L2 soundtrack and L2 subtitles (originally meant as support for hearing-impaired native speakers of the soundtrack language). Mitterer and McQueen (2009) found that watching subtitled television with L1 subtitles actually *harmed* foreign language perception and oral production of new sounds. In contrast, L2 subtitles helped to produce more native-like pronunciation. The authors indicate lexical interference as an explanation of why L1 subtitles are counter-effective in the learning process. In the experiment, the informants were exposed to accents of English that they were unfamiliar with (Scottish and Australian accents). Children and young people today, watching different English language television programmes with native-language subtitles, are in much the same situation. They are exposed over time to many different accents of English; national variants (British, American, Australian, etc.) as well as regional, local and social class accents. This may clearly be confusing when they try to build their own phonological system in English, as native accents differ both on the phonetic and the phonological levels. It is not unnatural to assume that this mixed input is so complex and difficult to grasp, that the overall consequence may be that learners resort to their native-language sound patterns. In other words, they may resort to L1 sounds since they lack a single, stable phonetic norm in the L2 they are exposed to, with phonological implications.

This discussion may easily be linked to the idea of *destandardisation*, used by Stuart-Smith (2011) with reference to the language expressions found in modern media. Traditionally, media was considered to have a stabilising effect

on language expression due to a more normative attitude concerning 'correctness', for example as reflected in the term 'BBC English'. Today, the situation is much more complex, with a host of channels sending different types of programmes aimed at various sub-cultures with different language expressions and identities. So Norwegian children today are increasingly exposed to English through the media, but the overall exposure is much more complex and varied than before, which may be a complicating factor when learning English as an L2 through the media, not least when it comes to phonological acquisition.

Conclusion and suggestions for further research

The overall findings in this study from Norway clearly indicate that massive exposure to oral English, mostly through subtitled television programmes, in itself has little or no impact on children's basic phonological acquisition in English as a second or foreign language. Young learners who have been exposed to English through the media on a regular and often daily or near-daily basis, at an age where they are still particularly receptive to language acquisition, seem to have the same problems when it comes to sounds and contrasts as learners of English of earlier generations, whose exposure to the language was much more limited. Pronunciation constitutes a basic element in the learners' active language skills, i.e. skills that form part of their L2 language expression. On that basis, it is fair to say that the findings support earlier research on incidental language acquisition through media exposure carried out on comparable groups of informants in Holland and the Flemish-speaking part of Belgium. Media exposure alone may contribute to the learners' L2 acquisition when it comes to passive vocabulary and L2 understanding, but as a potential tool in the process of acquiring active skills in an L2, it seems that this exposure needs to be exploited in structured teaching programmes where the learners are in an active learning mode rather than a passive entertainment mode. Passive exposure to English, even if quite substantial, seems not to be sufficient in relation to phonological acquisition, even in the age group most receptive to language learning. It is therefore important that English teachers develop their phonological awareness and knowledge about L2 acquisition processes. More emphasis should perhaps be given to basic, interactive pronunciation teaching in primary school, and in teacher education programmes.

Nevertheless, some recent studies (Kuppens, 2010; Lefever, 2012) seem to suggest that other aspects of language production among young learners, for example grammatical and conversational skills, are positively influenced by media exposure to an L2, so more research is needed to explore how, and to what extent, different types of active skills are influenced through such exposure. In addition, it would be of interest to further investigate practices in the Norwegian English classroom with regard to pronunciation teaching, and the

impact these practices actually have on young learners' phonological acquisition.

References

- Archibald, J. (1998). *Second Language Phonology*. Amsterdam: John Benjamins B.V.
- Bird, S.A., & Williams, J.N. (2002). The effect of bimodal input on implicit and explicit memory: An investigation into the benefits of within-language subtitling. *Applied Psycholinguistics*, 23(4), 509-533.
- Bush, M.D. (2000). Digital Versatile Disc (DVD): The new medium for interactive video. *CALICO Journal*, 17(3), 453-474.
- Cook, V. (2013). *Second Language Learning and Language Teaching*. Abingdon: Routledge.
- Doughty, C.J., & Long, M.H. (Eds.). (2003). *Handbook of Second Language Acquisition*. Oxford: Blackwell.
- d'Ydewalle, G., & Van de Poel M. (1999). Incidental foreign-language acquisition by children watching subtitled television programs. *Journal of Psycholinguistic Research*, 28(3), 227-244.
- d'Ydewalle, G., & Van Rensbergen, J. (1989). Developmental studies in text-picture interactions in the perception of animated cartoons with text. In H. Mandel and J.R. Levin (Eds.), *Knowledge Acquisition from text and Pictures* (pp. 223-248). Amsterdam: Elsevier Science (North Holland).
- Grgurovic, M., & Hegelheimer, V. (2007). Help options and multimedia listening: Students' use of subtitles and the transcript. *Language Learning & Technology*, 11(1), 45-66.
- Handeland, A. (1987). Diagnostic test of English (RP) vowels in the speech of Norwegian students. *Språk og Språkundervisning (Language and Language Teaching)*, 2, 25-35.
- Herron, C., York, H., Corrie, C., & Cole, S.P. (2006). A comparison study of the effects of a story-based video instructional package versus a text-based instructional package in the intermediate-level foreign language classroom. *CALICO Journal*, 84(4), 281-307.
- Hestetræet, T.I. (1993). *She Phoned the Teddy Bear*. Unpublished *Hovedfag* thesis. Dept. of English, University of Bergen, Norway.
- Huang, H.C., & Eskey, D.E. (1999-2000). The effects of closed-captioned television on the listening comprehension of intermediate English as a foreign language (ESL) students. *Journal of Educational Technology Systems*, 28(1), 75-96.
- Hulstijn, J.H. (2003). Connectionist models of language processing and the training of listening skills with the aid of multimedia software. *Computer Assisted Language Learning*, 16(5), 413-425.
- Johannessen, S.H. (1979). *Two Norwegian-English Regional Accents - an auditory analysis of the pronunciation mistakes in some 1st term Oslo and Bergen students*. Unpublished *Hovedfag* thesis. Dept. of English, University of Bergen, Norway.
- Koolstra, C.M., & Beentjes, J.W.J. (1999). Children's vocabulary acquisition in a foreign language through watching subtitled television programs at home. *Educational Technology Research and Development*, 47(1), 51-60.
- Kramsch, C., & Andersen, R.W. (1999). Teaching text and context through multimedia. *Language Learning & Technology*, 2(2), 31-42.
- Kuppens, A H. (2010). Incidental foreign language acquisition from media exposure. *Learning, Media and Technology*, 35(1), 65-85.
- Lam, Y. (2000). Technophobia or technophilia? A preliminary look at why second language teachers do or do not use technology in their classrooms. *The Canadian Modern Language Review / La revue canadienne des language vivantes*, 56(3), 389-420.

- Lefever, S.C. (2012). Incidental foreign language learning in young children. In Hasselgreen, A., Drew, I., Sørheim, B. (Eds.), *The Young Language Learner* (pp. 87-99). Bergen: Fagbokforlaget.
- Lefever, S.C., Ragnarsdóttir, B.A., & Torfadóttir, A. (2006). *Enskukunnátta barna í 4. og 5. bekk grunnskólans – Hvað kunna þau?* (English skills of children in primary grades 4 and 5.) Reykjavík: Iceland University of Education Research Centre.
- Lenneberg, E.H. (1967). *Biological Foundations of Language*. New York, NY: John Wiley & Sons, Inc.
- Leppink, M. (2010). *The Role of Media Input in Second Language Acquisition*. MA Thesis, University of Utrecht, Holland.
- Læreplanverket for Kunnskapsløftet. (2006). Oslo: Utdanningsdirektoratet.
- Macaro, E. (2003). *Teaching and Learning a Second Language: A guide to recent research and its applications*. London: Continuum.
- Markham, P. (1989). The effects of captioned videotapes on the listening comprehension of beginning, intermediate and advanced ESL students. *Educational Technology*, 29(10), 38-41.
- Markham, P. (1999). Captioned videotapes and second-language listening word recognition. *Foreign Language Annals*, 32(3), 321-328.
- Mitterer, H., & McQueen J. (2009). Foreign subtitles help but native-language subtitles harm foreign speech perception. *PLoS ONE*, 4(11), 1-5.
- Nilsen, T.S. (1989). Do you /ønde stænd/? Technical Report. <http://eric.ed.gov/ERICWebPortal/contentdelivery/servlet/ERICServlet?accno=ED321585> (accessed 14.06.2011)¹
- Nilsen, T.S. (2002). *British and American English Pronunciation* (2nd edition). Oslo: Universitetsforlaget
- Nilsen, T.S., & Rugesæter, K.N. (2008). *Basic English Phonetics for Teachers*. Bergen: Fagbokforlaget.
- Pinter, A. (2011). *Children Learning Second Languages*. Basingstoke: Palgrave Macmillan.
- Piske, T., MacKay, I.R.A., & Flege, J.E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics*, 29(2), 191-215.
- Pezdek, K., & Hartman, E.F. (1983). Children's television viewing: Attention and comprehension of auditory versus visual information. *Child Development*, 54, 1015-1024.
- Pezdek, K., & Stevens, E. (1984). Children's memory for auditory and visual information on television. *Development Psychology*, 20(2), 212-218.
- Rugesæter, K.N. (1992). The mystery of the missing phonemes. *Språk og Språkundervisning (Language and Language Learning)*, 4, 16-22.
- Rugesæter, K.N. (1995). The missing phonemes revisited. *Språk og Språkundervisning (Language and Language Learning)*, 1, 3-8.
- Simensen, A.M. (1998). *Teaching a Foreign Language – principles and procedures*. Bergen: Fagbokforlaget.
- Stuart-Smith, J. (2011). The view from the couch: changing perspectives on the role of the television in changing language ideologies and use. In Kristiansen, T. and Coupland, N. (Eds.), *Standard Languages and Language Standards in a Changing Europe* (pp. 223-39). Oslo: Novus.
- Taylor, G. (2005). Perceived processing strategies of students watching captioned video. *Foreign Language Annals*, 38(3), 422-427.

¹ This link has been temporarily deactivated by the provider (ERIC), but will be reactivated soon. ERIC number ED321585 – pdf pending release.

- Tschirner, E. (2001). Language acquisition in the classroom: The role of digital video. *Computer Assisted Language Learning*, 14(3&4), 305-319.
- Torfadóttir, A. (1994). Könnun á orðaforða i ensku meðal ellefu ára nemenda i nokkrum grunnskólum vorið 1993 (A study of English vocabulary of eleven-year-old children in compulsory schools in 1993). *Málfriður*, 10(2), 19-21.
- Vanderplank, R. (2009). *Déjà vu?* A decade of research on language laboratories, television and video in language learning. *Language Teaching*, 43(1), 1-37.
- Wells, J.C. (1982). *Accents of English: An introduction*. Cambridge: Cambridge University Press.
- Weyers, J.R. (1999). The effect of authentic video on communicative competence. *The Modern Language Journal*, 83(3), 339-349.
- Williams, H., & Thorne, D. (2000). The value of teletext subtitling as a medium for language learning. *System*, 28(2), 217-228.
- Yuksel, D., & Tanriverdi, D. (2009). Effects of watching captioned movie clips on vocabulary development. *The Turkish Online Journal of Educational Technology*, 8(2), 48-54.