TO [DʊI] OR NOT TO [DʊI]

Pronunciation in second language and second dialect: the case of Danish actors learning Norwegian

Åshild Livsdatter Brinchmann Løvvig

LING4190 - Master's Thesis in Linguistics
Department of Linguistics and Scandinavian Studies
UNIVERSITY OF OSLO
Spring 2023

Abstract

The theory of Embodied Cognition (EC) has in the recent decades come forth as a highly interesting and promising tool for language learning and learning processing. The possibilities EC provides by lightening the cognitive load when processing new information is shown by a large body of research, and is one of the main inspirations of this thesis. The other main inspiration is acting, and its possibilities to be used as a useful tool and an object of study within linguistics. The goal of this MA thesis is to explore the use of embodied action as a tool in pronunciation improvement, in the context of using acting as an embodied action in aiding pronunciation improvement. Over the course of this thesis, various types of data has been gathered to be analysed, mainly speech recording sessions on the L2 learners involved, and rating sessions with native speakers of said language, as well as the use of questionnaire and an interview. The participants were three Danish students in Norway studying theatre acting, and were receiving Norwegian classes in their school programme. The students went through three recording sessions done over the course of one school semester, which included reading practice sentences that contained phonemes challenging for Danish speakers, inspired by their language classes, a questionnaire and an interview per session. The recorded speaking of the practice sentences were analysed in Praat for accuracy rating. Native speech by three native Norwegian speakers was also recorded, as comparison data. All the recording data was used in a rating session done by another three native Norwegians judging the accentedness and comprehensibility. The results show an improvement of overall pronunciation accuracy, and ratings on accentedness and comprehensibility rating improved over time. Some phonemes had little improvement, mostly due to already being high in accuracy rate from the first session, while some phonemes, particularly those with initially lower accuracy rate, improved over the three sessions. What the data did not show was that action had a positive effect in lightening the cognitive load, meaning that action possibly ended up being cognitively taxing for the participants in spontaneous speech when acting. The results nonetheless show that the use of action does not inhibit improving pronunciation, and it is inconclusive whether adding action to speech may or may not help learning L2 speech over time.

Acknowledgements

This thesis would not have been possible for me to write without the immense help from several people.

First I need to thank my supervisors, the endlessly helpful Peng Li, who aided in more ways than I could expect, pointing me in all the correct directions. He gave me materials, taught me how to collect, store, and analyze data. He kept my spirits up with his light mood when I feared I was lost in all the data and scientific papers. He answered my emails with encouraging words, even though I could be both panicked, somewhat frantic, and at times, uncertain if I ever would finish my thesis. I also want to thank my other supervisor Haley de Korne, for aiding me (and Peng) in formalities, lending me the recorder, and giving amazingly helpful feedback during my writing. Academic English is, I (re)discovered, very difficult, and your help with vocabulary suggestions and proofreading helped me sound way more competent than I find myself to be. I believe my sometimes casual tone in this thesis could seem very amateurish and un-academic, and I thank you for the patience and grace in feedback on those parts especially.

I want to thank the three students, Thyra, Margret and Alexandrine (pseudonyms) who contributed to almost all the data I used in this thesis. Your willingness to lend me your time, your patience with me, and your bravery in acting, living, and breathing a similar, but yet foreign language, is highly admirable. You have come extremely far, and I am deeply impressed by your progression. I truly hope the Norwegian theatre community will realize and recognize your hard work and will include you in many productions over the years. A big thank you to the students' two language teachers and the educational art institution they worked at, who provided me with a lot of information, let me be a part of the language classes, and gave me materials I needed.

I want to thank all of the teachers I've had throughout the years at UiO who encouraged me, sparked my interest even further (so I could further annoy all my non-linguist friends with new facts and theories), and inspire me in academia, particularly Pritty Patel-Grosz, Patrick Grosz, Åshild Næss, Pia Lane, Andreas Sveen, Timo Roettger, Pernille Hansen and Nicholas Elwyn Allott (and his cats), and other names that might have slipped my mind.

I also want to give a great thank you to Lene, Maria K, and Katinka for providing your native (beautiful) voices. Thanks to Joachim, Aili, and Maria H for being my (very patient) raters. An

extra thank you to Joachim who also proofread for grammar and spelling mistakes. A big thank you to my mother, who also did proofreading, and provided me with much needed academic advice, and always helps calm me down. Thank you to my father and my little sister, for being my steady family. Thank you to my friends who had beer and coffee with me to ease my mind, and to all who said I could do this. Thank you to Sofia for sending me encouraging words and pictures, and one poem.

Lastly, I want to express my deepest gratitude to my boyfriend Stian Ofrim, who formatted the graphs, for being my conversational partner when I needed to purge thoughts and anxieties, and for being my personal chef, caretaker and *bautastein*, for your endless support and encouragement (and patience). You complete me.

Contents

1. Introduction

When it comes to learning a foreign language, there are many aspects to acquire. To learn a new language entails a new vocabulary and a new grammar (including linguistic characteristics like case, articles, pluralization, honorifics, and grammatical gender). Another important part of language acquisition is the pronunciation, which is inevitably impacted by the accent you bring into the language you are learning.

There might not be such a thing as a "perfect accent" in any language. It is not even necessary to sound exactly like a native to be understood, as we have the ability to understand a certain deviation from our frame of reference for a native accent (Munro & Derwing, 1995, 2020; Galante & Thomson, 2017). To attempt to "get rid of" one's own accent is a personal, aesthetic choice, often motivated by social and political pressures. It is nonetheless of interest to the field of language learning and acquisition to examine the many why's and how's of pronunciation. There are also certain fields and disciplines of work interested in the details of accent, dialect and speech particularities, and how to manipulate, change, and practice their speech. The particular professional field of interest in this study is acting.

Actors are no strangers to learning dialects or taking on an accent for a particular role. There are many famous examples of this within film, from the impressive work of Idris Elba taking in a particular Baltimorean AAVE dialect in *The Wire*, or Jennifer Lawrence speaking Appalachian Mountain English in *Winter's Bone*, to the less fortunate examples like Keanu Reeves attempt at British English in *Bram Stoker's Dracula*, or Dick van Dyke trying for a Cockney accent in *Mary Poppins*. The use of accent training and accent change is just as present within the theatre, and the educational frames of acting, which forms the context for this study.

In the past few years, more and more students from Norway's neighbouring countries and close linguistic friends, Sweden and Denmark, have come to Norway to take their theatrical education. This has created opportunities to give these particular students language classes, to learn to speak Norwegian for the Norwegian stage.

In this thesis, the focus has been on three Danish students studying acting in Norway. Danish and Norwegian are, politically at least, different languages. Linguistically, it could be debated if they are in fact dialects of the same language (Torp, 2006). Nevertheless, the great similarities

between them, allow for a unique exploration of learning a language where grammar learning takes the backseat, and pronunciation, accent, and comprehensibility are the main areas of focus for the language-learner.

Aims of the study

This study aims to explore pronunciation learning through the theoretical lens of Embodied Cognition (hereafter EC). The expanding theory of EC has reached beyond the field of cognition, and is of interest for linguistic research (Foglia & Wilson, 2013; Shapiro & Stolz, 2019; Wilson, 2002). Although research on EC started decades ago, there seems to be a need for further studies in a field that, to put it metaphorically, has left its infancy, but is barely walking on two legs, and is far from running. As examined further in chapter 2, the theory of EC refers to the notion that mind and body function as a duality in cognition, and are equally essential in how we can facilitate processing of information and learning. The examination of the nature of the body and mind is a fascinating field deserving of further exploration in many areas of science. I argue in this thesis that the duality and cooperation between mind and body is equally important when learning language by putting this thesis in the context of previously written literature that focus on language learning and various methods that tie together with EC.

The study of EC can also bring together various different fields and professions. In this case, I have looked to the world of acting and theatre, as well as the field of linguistics and language learning, two seemingly very different worlds, except for their quite obvious shared interest: language, and the detailed study of communication.

Within the world of theatre, the ability to change your own voice to a certain extent, whether it is to make it sound a different age or of a different class, or simply to make it sound like a different dialect, is a desired skill for many, including those who act and perform outside of their native language. While *getting rid of* an accent completely is by no means the only valid thing to do when learning a new language, the close study of how to train our perceptions and pronunciation skills can be of interest to researchers of the field of language learning, phonology, even socio-, neuro- and psycholinguistics, as well as actors learning a new accent for a role.

The role of accent or pronunciation is thus important in both theatre and linguistics. In this study, I use the term accent to refer to how different an L2 speaker's pronunciation is from a target

variety (Munro & Derwing, 1995, 94), and accentedness refers to the perceived differences in pronunciation as compared with a local variety (Derwing & Munro, 2015, 14).

Pronunciation refers to the realized speech produced by a speaker. Another key term is comprehensibility, which refers to the perceived degree of difficulty experienced by the listener in understanding speech (Derwing & Munro, 2015, 14). All of these concepts are discussed in more detail in chapter 2.

My motivation for this study is both scientific and personal. I myself have an acting and theatre background, with a degree and years of experience within different kinds of theatre practice. It is a passion and interest that still lies within, and I believe academia and the arts have a lot of potential in working together, by merging and mixing the disciplines (see also studies of dance and syntax, e.g. Patel-Grosz et al., 2018, 2022).

In this study I want to explore Embodied Cognition in relation to acting, and how it impacts language learning, specifically pronunciation training and accentedness. Secondly, I want to explore exactly how pronunciation training can be aided by an EC-based theatre and acting dialect training. There lies potential in finding answers to these questions. It is of interest to find effective methods of language learning within fields in SLA (second language acquisition), SLT (second language teaching), and cognitive linguistics, and actors working closely with speech and text may benefit from the awareness on the duality of the body and the cognitive, and how this may aid in dialect/accent training, and possibly other disciplines within the acting craft.

To explore these questions, I have been an observer of native Danish students learning Norwegian through pronunciation practice at an educational art institution in Norway. Their training has incorporated theatre techniques such as adding meaning, subtext, and inner goals to the practice sentences used. After observing their training and discussing their learning with their instructors, I designed a mixed methods quasi-experimental case study, as explained further in chapter 3, with the aim of understanding whether the theatre techniques lead to better or more confident results in their pronunciation of Norwegian.

In this study, I chose to focus on a specific kind of *embodiment*, or *bodily movement*. Indeed, mouth movement and articulatory movement is bodily movement, but the interest here lies in bodily movements other than the necessary articulatory movements, since these students

learning Norwegian had already been learning for a while, enough to be considered intermediate speakers (there is also the natural advantage of the two languages being very similar, and some previous exposure is almost sure to have happened, considering the Scandinavian neighbouring relationship). In this thesis I am arguing that the body participates in the L2 pronunciation online, in that we move when we speak, which would ideally lead to partly offloading the semantic information onto the bodily movements, which opens the possibility for lightening the cognitive load. Lightening of the cognitive load will then give room for cognitive resources to focus on pronunciation.

Acting (with all its subtle variations) involves utilising both body and mind, and the ability to place meaning and emotional gravity on learned (or improvised) text (Merlin, 2007). It is a complex, multimodal process of combining the physical with the mental processes of getting into a role and adding goals, emotions and meaning to a scene. To add these emotions and meaning to a text or speech also adds motions and actions to the body, making the acting into a full performance of voice and body. These "actions" are in focus in this thesis.

The participants performed actions along with the practice sentences they were given, through adding elements that encouraged some form of acting. In essence, *acting* and *action* should be treated as one and the same in this thesis; to trigger acting is to trigger action. Every time there is no acting/action present, this will be referred to as *non-action*. Together these will make up two conditions: action and non-action.

In this paper, I seek to answer the following questions:

1. Is there a substantial or noticeable effect of the use of action vs. non-action on the pronunciation accuracy of nonnative sounds?
2. Is there a substantial or noticeable improvement with the use of action vs. non-action with overall oral proficiency on the measures of comprehensibility and accentedness?

Structure of the study

Following the introduction, in chapter 2 I will discuss the theoretical and empirical background of this thesis. Chapter 2 is divided into seven sub-chapters. First, I examine L2 and language acquisition (2.1), and accentedness and comprehensibility focused research (2.2). I then move on

to a section on acting interpreted as a *bodymind* discipline and what this entails for the author (2.3). Further, in chapter 2, I will introduce the foundation and main theory behind the thesis, Embodied Cognition (EC) (2.4), before moving on to different empirical studies related to EC and language learning (2.5, and its subsections), and the use of drama and acting in relation to language learning (2.6). Finally, I will move into the phonetic differences between Danish and Norwegian and how that affects learning as an L2 speaker (2.7).

Chapter 3 presents the methodology I have employed, including the research design (3.1), a description of how the participants were learning Norwegian (3.2), and the data collection (3.3).

Chapter 4 will present the results and analysis of the thesis, divided into five sub-chapters. The sub-chapters in chapter 4 will present the results of accuracy rate of each participant (4.1), accentedness and comprehensibility rating results (4.2), results and analysis from the questionnaire (4.3) and the interview (4.4), before summarizing the results (4.5).

Chapter 5 discusses the results of the previous chapter with a brief summary (5.1), then I answer the research questions based on the results (5.2). I discuss the importance and impact action can have on the relevant measures in the thesis (5.3), talk about EC in language teaching for this particular context (5.4), mention possible future research (5.5), and the limitations of this study (5.6). Chapter 6 presents the conclusion to the thesis.

2 Theoretical and empirical background

This chapter introduces the theoretical background and previous research on language learning, accentedness and comprehensibility, the role of drama in the classroom, and the introduction of the term Embodied Cognition, as well as some research on acting and how it can relate to Embodied Cognition.

2.1 L2, language acquisition, and language learning

A part of learning a new language is learning a new phonetic system (in addition to syntax, new semantics, slang, cultural context etc.). For many, this might be the hardest part of learning a

new language. Learning to pronounce like the native speakers can be compared to *rearranging* one's own mouth, to place the tongue and lips in ways you've never done before.

Foreign accents, in all languages, is a form of divergence from the phonetic norms of the given language, and it may have undesirable consequences for the L2 learners, like being misunderstood or being difficult to understand by native speakers. It can also create negative self-evaluation or evoke negative group stereotypes (Lambert et al., 1960; Giles, 1970).

The causes of foreign accent can be many, from the lack of neural plasticity due to aging (Penfield, 1965; Lenneberg, 1967), inaccurate perception of L2 sounds (Flege, 1992a, sited in Flege 1995; Flege, 1992b), to lack of motivation, input or incorrect habits (in relation to the learning of L2) (Flege, 1988b, sited in Flege, 1995). It has been observed that the further out in life one begins to learn a language, the stronger the accent becomes (Flege, Munro & McKay, 1995).

This seems to be what many of us experience. Adults learning a new language usually seem to have difficulties exercising their pronunciation to a point where they sound exactly native. For many, the most important thing is to be understood, and not everyone wants to get rid of their accent, which naturally is a perfectly valid choice, and can add diversity to a language (Farrell & Martin, 2009). But some degree of accent is sensible to practice away, considering that heavily accented speech may cause more effort to be understood. We practice to further round the lips, practice aspiration or de-aspiration, learn tones, lower the back of the tongue or snap with the tongue to make a click, all in repetition, to bring us closer to our target sounds. Flege (1995) says that "bilinguals" are prone to hear sounds in L2 through the "grid" of their L1 phonology (here, "bilinguals" seems to refer to only monolinguals learning a second language, while naturally bi/multilingualism from early adolescence is highly common in the world), which ensures that non-natives perceive some L2 sounds differently from native speakers (Flege, 1995, 237).

SLA in practice as second language learning (SLT) has had a shifting focus on different methods in pedagogy, with a focus on usage-based learning, going from grammar and translation to teaching spoken language, and then starting to focus on acknowledging learners' needs, learning styles and motivation to achieve communicative competence. There has also been a shift from achieving native-like proficiency in English, to acknowledging World Englishes (Savvidou, 2020; Farrell & Martin, 2009; Crystal, 2003), discussing the problematic connotations of the

term “Standard English”, and how accents and different grammars within English (ex. Singaporean English and its absence of possessive inflections and use of particles) contributes in diversifying English as a language without a standard, while also considering the issue of teaching English as a foreign language and how to balance the language teaching with informing of the varieties (Farrell & Martin, 2009). This could naturally be applied to other languages as well, particularly majority languages in any given country with a substantial immigrant population. Cook (1999) also writes that an L2 learner can possibly become native-like in their speech, but cannot, by definition, ever become a native speaker “without being reborn” (Cook, 1999, 187). But how native-like can an L2 speaker become?

Flege and his colleagues developed a speech learning model, or SLM, that focused on experienced learners of an L2, and how and why their age-related limits on abilities in pronouncing L2 correctly (native-like) functions, claiming that L2 sounds will be inaccurate when lacking accurate perceptual “targets” to guide the sensorimotor learning of said L2 sounds. This included five postulates and seven hypotheses to be reviewed and discussed (1995, 237-238), which I won’t discuss in detail here.

The SLM theory was later on revised into the simply titled SLM-r; speech learning model revised (Flege & Bohn, 2021). While SLM had a more apparent focus on the Critical Period (CP), the thought that diminished cerebral plasticity leads to reduced ability to learn L2, with differences between groups of individuals learning before and after (the supposed important) CP, SLM-r focused more on “how phonetic systems of individuals reorganize over the life-span in response to the phonetic input received during naturalistic L2 learning” (2021, 3), while also accounting for differences between early and late learners. This meant that SLM-r abandoned the focus on the highly experienced L2 learners and how they could “master” the L2 sounds.

Another difference was that while SLM proposed that there would be an upper limit on L2 production accuracy based on the accuracy of L2 segmental perception, the SLM-r proposed that L2 segmental production and perception coevolve equally (2021, 28-29). According to Flege, “Categories formed for L2 sounds are defined by the statistical properties of input distributions” (2021, 40). He writes that SLM-r also maintains that L2 distributional learning echoes L2 distributional learning. This means that, since the distributional learning is slow in L1 acquisition, the same will happen in L2 learning.

Although it is widely considered that learning L2 earlier rather than later has clear benefits, it only works in the long run. Where adults out-perform younger learners in the early stages of naturalistic L2 acquisition (focus on exposure), the adult-child differences shift over time, depending somewhat on the exposure, where younger learners have the advantage over time (Flege & Bohn, 2021; Jia, Strange, Wu, Collado, & Guan, 2006).

Although some apply the "downfall" of learning skills of adults over time to age-related cognitive changes (DeKeyser & Larson-Hall, 2005), SLM-r claims that all learners of an L2, regardless of age, uses the same mechanisms and processes that children exploit when learning their L1. It also claims that we do not lose cognitive capacity to learn speech over time, but that those mechanisms and processes do not yield the same results when applied to learning sounds in an L2 as they did in learning L1 (2021, 23). And so, L2-sounds get linked to the L1 phonetic inventory , which can interfere with the new L2 phonetic inventory. In other words, there is a perceptual bias induced by the L1 phonetic system (2021, 24). This should mean that adult learners actually *can* physically, or cognitively (see subsection 2.4), learn better, and even excellent, pronunciation later than the assumed critical period. Flege also discredits Lenneberg's (1967) CP hypothesis as misleading and incomplete (2021, 24), for example that accents are indeed present even in pre-critical period learners of L2, and that learning through exercising great effort and labour to the target language is quite subjective to the learners of different types (like foreign language classroom learning versus immersion through immigration).

From this point of view, even if there is such a thing as a critical period, it is more fluid than initially thought, and it gives no hard line for when and how well learners can develop their skills. However, one notice Flege writes, is that he claims that it is virtually impossible for L2 learners to be identical in production and perception to "mature monolingual" native speakers, and that it is therefore not of theoretical interest to be able to determine if an L2 learner is indistinguishable from native speakers (2021, 25).

Generally, the focus is seemingly on the "ideal" adult monolingual speaker, and Flege even claims that being bilingual may prevent the so-called mastery of L2 sounds (2021, 26). Bilingualism is presumed here to refer to learning a second language later in life than at infancy, which of course many do. Does this mean that learning a language later in life than the very earliest years cannot possibly even become perfect? I have anecdotal experiences that this is not

so, with people of many ages and nationalities learning Norwegian to a degree of sounding native. But we cannot lean on the anecdotal. Flege argues that the formation of a new L2 phonetic category sound is, among other things, dependent on the quantity and quality of L2 input obtained for the sound in meaningful conversations. This could very well mean that everyday speech involving meaningful and engaging content would be considered both of better quality and quantity than classroom speech learning, with practice sentences taken out of context, without a natural speech flow and social engagement (Flege & Bohn, 2021).

Lastly, to elaborate on Flege's thoughts on the interference of L1 on L2, he writes that the source of pronunciation difficulties and accent is caused by three factors:

- That L1 sounds "substitute" L2 sounds, and the L2 sounds become automatically linked to sounds in the L1 phonetic inventory
- That preexisting L1 phonetic categories may then interfere with, or even block, the L2 sounds as a formation of new phonetic categories.
- That the input of L2 sounds differs from the input that monolingual native speakers of the target L2 receive when learning the same sounds. Instead of forming a new sound category for an L2 sound that differs phonetically from the closest L1 sound, a composite L1-L2 phonetic category will develop that is based on phonetic input from *both* languages (2021, 23-24).

The interference of L1 sound categories with L2 sound categories, and the notion of *substitution*, are similar to the concept of *transfer*, which has been used to explain the possible role of L1 influence on the interpretation of prosodic variations (De Marco, 2020). L1 transfer plays an important role in interpreting paralinguistic intonational meaning in L2 (2020, 8). One example being that Russians and Persians expressing themselves in Italian (B2 level) expressing emotional utterances in their L2 Italian in the likeness of their L1, differing in their emotional productions from the L1 Italian speakers (De Marco, Sorianello & Paone, 2016 in De Marco 2020), essentially meaning that L1 emotional speech transfers to their L2, possibly resulting in less effective communication. This is also said to be a possible source to misunderstandings and miscommunications with native speakers when produced by L2 speakers (2020, 11).

De Marco (2020) has written an important addition to SLA on the difficulties of expression and perception of emotional states in a foreign language, citing the expression of emotional states in L2 requiring full control of the prosodic resources that contribute to their realization (2020, 5).

According to De Marco, training focusing on expressing a wide variety of emotions, as well as video dubbing have proven to be useful tools in improving production performance and perception of prosodic patterns of emotional communication. These “hidden” and “subtle” communicative dimensions are no less essential in communication, giving form and meaning to our words, like prosody and facial expressions, gestures and postures do (2020, 5). Different meanings and nuances are realized through the human voice’s prosodic modulations. These prosodic modulations, representing suprasegmental entities, “merge with the segmental characteristics typical of each language” (2020, 6), which I would argue is what makes a language “whole”, and is, as mentioned, often a dimension of language that may separate L1 speakers from L2 speakers, and could easily be under-communicated as a factor to L2 speakers. It is important to recognise this part of language learning and language acquisition as well when talking about mastering a language, and becoming “fluent” (discussion of the definition of this term will ensue later), even though these variations are easily overlooked.

Considering some languages are closer than one another, it would be expected that two highly similar languages would have an easier time skipping the issue of transfer, explained by researchers finding a higher percentage of emotion identification when having in-group advantage and linguistic closeness showing better decoding process of emotional expression (Scherer, Banse & Wallbott, 2001; De Marco, 2020). This would then contribute to an increase in native-like speech. However, there are still many segmental and suprasegmental traits of a foreign language, no matter how similar, that will prove to be challenging to an L2 speaker, especially if there are traits not existing in the native language of the L2 speaker.

In this thesis, my focus is on learners that already speak two or three- languages, and that are learning a language that is considered to be rather close to their native language, close enough that in some circles, they might be considered dialects of one another (Kristoffersen, 2000; Torp, 2005). So far there seems to be little to no research done on natives of a language learning a (very) closely related language later in life. Learning such a closely related language gives a lot of room for naturalistic learning, with the possibility of input from everyday conversations and society around the learners. In learning such a similar language, learning a new vocabulary can take the passenger seat, and the articulation and pronunciation can receive most of the practice and focus. Similarly, there is minimal amounts of studies done on acting with a linguistic perspective. I hope to contribute to this particular narrow type of study.

2.2 Accentedness and comprehensibility

Accentedness is mentioned as a correlation of oral proficiency by Galante & Thomson (2017): "fluency, comprehensibility, and accent [...] are widely accepted as important correlates of oral proficiency" (2017, 116). The researchers base their definition on Munro and Derwing's (1995) distinction between comprehensibility and accent, where comprehensibility refers to listeners' subjective sense of how easy a stretch of speech is to understand (a processing perspective), whereas accent refers to how different an L2 speaker's pronunciation is from a target variety, that is, the given language and/or dialect they're learning. (Munro and Derwing, 1995; Galante & Thomson, 2017). Although it is not a focus of this study, it is worth mentioning the distinction between comprehensibility and intelligibility: Comprehensibility is about how easy or hard it is for a listener to understand speech, while intelligibility refers to how much a listener actually understands of a speaker's *intended* utterance (Derwing & Munro, 2009; Galante & Thomson, 2017).

Naturally, one wants to be comprehended and to be intelligible, and the ability to understand certain accentedness might differ from person to person. Another important distinction is between accent and accentedness. Accentedness is defined by Munro & Derwing (2015) as the perceived differences in pronunciation as compared with a local variety (2015, 14). They write that studies focusing only on accentedness, follow the nativeness principle, emphasizing promotion of native or native-like speech among L2 learners, might not offer much benefit for classroom teaching because of the somewhat lack of focus on communicative effectiveness. This is arguably true to some extent. Focus on accentedness alone will not help learners excel in their language learning, particularly not if the learners are still somewhat inexperienced in the target language. However, many studies focus on overall pronunciation, comprehensibility, fluency, flow, enjoyment, confidence, and memory as well as accentedness (studies which I will get to in subsection 2.4 and 2.5). A focus on accentedness might also be of interest to learners who already know the target language well, either through long experience, or because the target language is highly similar to the learners' native language.

According to Munro & Derwing, accentedness and comprehensibility (as well as intelligibility) are somewhat independent of each other. Research tells us that heavily accented speech can score high in comprehensibility (Munro & Derwing, 1995, 2020). Listeners have also tended to assign more negative scores when rating accent than comprehensibility, and they could even

transcribe sentences perfectly despite the heavy accent (2020, 17). This research shows empirical evidence that a heavy accent does not entail reduced comprehensibility for listeners.

Munro & Derwing writes of this approach to L2 speech evaluation as a possible useful tool in investigating what methods may be more beneficial when teaching pronunciation to ESL learners (1995, 286). This could possibly be true for other language learners as well. A listener would likely judge accentedness on the extent to which the pronunciation deviated from what native speech would sound like (2020, 18). So the less an accent deviates from native speech, the better the accentedness score (and vice versa). Although accentedness and comprehensibility are closely tied, listeners would likely be able to judge comprehensibility regardless of how high or low they would rate the accentedness, based on the evidence found so far.

It is worth taking into consideration that individual listeners have somewhat individual differences in perception to non-native speech, like in the relationship between accentedness and comprehensibility (2020, 18). For Norwegian listeners of L2 learners, some could judge a mispronunciation of a vowel heavier than others, while some would judge a wrong toneme as more of a serious error than others. All this considered, variations in opinion and rating is to be expected, and it is up to the researcher how many steps they should take to minimize possible differences in perception. But there is only so much one researcher can do to create a fixed picture for various L1 speakers of a target language.

As priorly mentioned, the focus on accentedness alone could in many cases not benefit classroom teaching of language. But there are other ways than classical classroom teaching, like involving a more engaging teaching method, like including activities and role plays. As I will present further in this thesis, the use of role play and drama in classrooms have proven to be effective in several ways.

In addition, an L2 speaker might have mastered grammar, structure and fluency to a native-like level. On notes on the definition of fluency, it usually means general proficiency in lay terms, involving a certain kind of speed, the automatic use of language in real time, and factors such as flow and continuity (See Galante & Thomson, 2017 for suggestions to how fluency has been defined).

Norwegian and Danish are two languages that are close in both pronunciation and grammar, especially the latter (Torp, 2006; Kristoffersen, 2000). The languages still have some very apparent differences in pronunciation, however. For a Norwegian to learn Danish, or a Dane to learn Norwegian, the focus on grammar would be minimal because of the similarities, and as a study it would be a somewhat rare case of learning a language where learning to adapt the native accent is the main focus. The quite apparent phonemic differences will be discussed later in this chapter (2.7).

It could also be assumed that learning the other language would take considerably shorter time than one person learning an L2 which would be topologically very different, again considering the similarities of grammar. Although the necessity for a Norwegian or Dane to learn the other's language is not immediate or of great need (although some linguistic adapting to one another is not uncommon to be heard among the neighbouring nationalities), the case of Danish acting students in Norway learning Norwegian includes an interest to learn the language as close to the native speech as possible, mainly because of job opportunities, among other things .

Studying language learning in such a context offers the opportunity to study the progression on accentedness and comprehensibility on fairly experienced learners of an L2, in this case, Danish learning Norwegian. The next section will introduce research on acting, to introduce the reader to some theory on acting and the use of mind and body, including voice, and how these play together as a whole for the actor's craft.

2.3 Acting, theatre and the "Bodymind"

Here I will present some writing on the actor's experience and the embodied modes of experience for actors. The essay *Toward a Phenomenological Model of the Actor's Embodied Modes of Experience* by Zarrilli (2004) is meant as a contribution to phenomenological studies of embodiment, but can be used as a contributor to the field of EC as well, especially for the working actor and their modes of embodiment (2004, 654).

Like with cognitivism, Western philosophers came to the conclusion that the body is a physical object, like other material objects, in the beginning of the seventeenth century. A series of books in the 1960s by Maurice Merleau-Ponty challenged this, in calling focus to the "lived body", instead of an object body, and the embodied experience. Instead of being a representable object,

the body should be seen as an experienced phenomenon. This is, however, essentially a rejection of the mind-body dualism (2004, 654-655), even though there was a distancing from the objectification of the body as an almost optional vessel of sorts. This later was challenged, when Drew Leder added a "post-Merleau-Ponty" account on the "corporeal absence", and in this case, Leder argues that it is precisely our experience of the absent body that is the foundation of Descartes' body-mind dualism (Leder, 1990 in Zarrilli, 2004, 656). Zarrilli further examines this, but focusing on the extra-daily modes of embodiment, rather than only on Leder's everyday modes of embodiment. Zarrilli then introduces the aesthetic "inner" bodymind and an aesthetic "outer" body, as an form of embodiment. The "aesthetic" refers to the non-ordinariness, that the embodiment takes place over time, and that the experience of body and mind separation turns to a subtle engagement (2004, 661).

In short, to Leder, the "Absent body" is our lack of attention to our body and then the lack of "experience" it receives. This is exemplified by situations like reading a book, being lost in thought, then ignoring physical sensations. The same, he says, goes for engaging in sport, where muscles flex and the slightest movements of opponents create immediate reactions, where still the full attention goes towards the game and the opponent (essentially, the other body/bodies in front of you), and attention falls away from embodiment (Leder, 1990 in Zarrilli, 2004, 656). Naturally, when practicing a new skill, like a sport or learning to change a tire, our attention towards our body is quite different, and our attention dwells more in how we move and use our body. Skill acquisition, says Zarrilli, is at first extrinsic, before becoming more intuitive, or *in*trinsic, and the "bodymind" hence adjusts to the new movements of different kinds, like arms and legs in martial art, or precisely our tongue and lips in a new language (2004, 658-659). Zarrilli refers to this as the body essentially "disappearing". This, I would argue, could also be called a form for "merging" of body and mind, where the now acquired knowledge steers how the body moves and flows, or how the tongue and lips follow the knowledge of a language (of course, a language is far more than just pronunciation). Like Zarrilli says, one goes from acting "*to* the skill" and ends up, after practice, acting "*from* the skill" (2004, 658-659).

In acting as a profession and a discipline, the body itself is of the central importance, even though acting often is all about "pretending" to be somebody else, or even some*thing* else, put in very simple terms. In theatre, the body is equally important to all forms of acting and theatre (even in such a play Samuel Beckett's short monologue play "Not I", in which the only character is a pair of female lips illuminated on stage, speaking at a ferocious pace). The action of acting,

and similarly, other extra-daily activities (yoga, martial arts) over a long-term period engagement, are part of what Zarrilli called the “inner bodymind”, mentioned earlier. It is embodied practice which engages both the (physical) body and the mind, and attunes them to levels of experience and awareness (2004, 661). He even calls this process of cultivation a “dialectal engagement of body-in-mind and mind-in-body” (2004, 661). One could essentially say Zarrilli expresses a way of thinking about the cooperation between body and mind, in the sense of acting and theatre practice, where one can both work “outwards” and “inwards”, attention shifting depending on what practice or activity one does.

Within the context of theatre art, and extra-daily physical activities and disciplines like yoga, meditation, martial arts, even poetry reading/writing and song, the idea of an inner bodymind brings forth an element of awareness to how the mind and body works together when exercising both incorporated skills and new skills still being acquired. It should be no controversial stance that both elements are equally important for such practices and activities, and to actively understand them can give an increased understanding of one's own practice. There is certainly some difference in the way linguistic academics and teachers and masters within the art of theatre (in all its forms) talks about embodiment, and how we use the brain and body, in what context, and in relation to whom. I include this essay by Zarrilli to add some perspective of the field of which this study will dive into. Theatre and acting is certainly more rooted in personal reflection and viewpoints than academics can be, but although the two disciplines might seem hard to combine in an academic study, I attempt to give it a try.

2.4 Embodied Cognition

The idea of Embodied Cognition (EC) first entered the field of Cognitive Linguistics as early as in 1975, which from then on we saw the dawn of a neural theory of thought and language, or NTTL. George Lakoff explains NTTL that thought is physical, that thought is made meaningful by the ways our functional neural circuitry is connected to our bodies, and that the abstract ideas and language of our minds are equally embodied (Lakoff, 2012, 773). In this chapter I will present in detail various research on the field of EC. Different researchers and linguists have attempted to explain the theory that mind and body are interconnected in a multimodal, sensory-motoric unity, and further on, we will look into how this theory has been put to use in how we learn language.

Cognition is affected by the interaction with people and the environment through the body, and the processing of all this information can be facilitated through our embodiment and different levels of physical aid. This can be anything from using your fingers to count or point to or towards the relevant elements (e.g. to use the fingers to figure out which number in line pertains to which month), to turn or tilt our heads to either see where an unknown sound comes from, or to read the spine of a book, to write down things we have to remember, to use visual aid to explain and illustrate different kinds of concepts to others (e.g. a Powerpoint in a presentation, small-scale models of real constructions, or show test patches of fabric to touch and feel) and so on. We use our bodies' movements, and interact with our environment to aid us in reducing cognitive load, making information easier to take in, process, and remember for longer amounts of time. From this, I will present some research rooted in EC. This will be the fundamental frame of reference for this study.

EC can be seen as a contrasting view to cartesian dualism. In cartesian dualism, mind and body are fundamentally separate, and has influenced various strands of cognitive sciences for a long time. EC goes the opposite way, and sees mind and body as closer related, where sensory and motor functions are seen as a part of our cognition, with cognitive skills as "the product of a dynamic interplay between neural and non-neural processes" (Foglia & Wilson, 2013, 1). In this sense, the body of the agent/individual, cognition, and real-life contexts are in one feedback-driven ensemble, all giving back to our mental functioning (2013, 1). The body itself doesn't seem to necessarily be directly linked to cognition, but can have an "implied role in reenactments of experience in the brain's modality-specific systems for perception and action" (2013, 319)

EC argues that mental activity is body-based. So what EC is *not*, is a *di*sembodied mind, with symbolic representations that are autonomous from the sensorimotor system, and are amodal, abstract and arbitrary (2013, 2), where the bodily actions and motor system are separate and independent from cognition. In a radical sense, one could almost deduct that traditional cognitive views see the brain as one entity separated from the physical, "real" world, and that all mental representations are cognitive only on their own. As an exaggeration, in this world, a floating brain in a glass jar filled with alien-like bubbles, could be possible.

EC seeks to challenge these notions, where the cognitive and embodiment are closer connected, and possibly inseparable. Differences in one translates to differences in the other, and,

interestingly put by Foglia & Wilson: “algorithms that constitute cognition sometimes reflect the peculiarities of the physical body” (2013, 320). Direct engagement with sensory input also does not necessarily need to be present for our sensorimotor functions to activate, though they definitely can be. Foglia and Wilson cite imagery, planning and remembering as non-direct engagement with sensory input and behavioural output.

Further, there are several kinds of observations about our behaviour that Foglia and Wilson mention as motivations for the embodiment thesis. I will list them here:

- Gesturing:

Gesturing while speaking affects interpersonal communication and language processing, it can clear up information to the listener-viewer and picture the speaker’s mental representations. Similarly, we use our fingers to help counting and understand number concepts, or point to something to manipulate focus of attention to something.

- Visual perception and bodily movement:

We move and we correct our bodily movements in order to perceive better or differently, like that we have to tilt our heads to properly read the spine of a book, tilt and turn any given object to get a view of the whole object, or move closer or further away from something to change the perspective and information we’re given of our environment or the object we’re viewing, or to turn around in response to an unexpected noise (2013).

- Mirror neurons:

Our mirror neurons activate during observation or understanding of an action performed by others, and when carrying out the same action with our own bodies. There have been suggestions that yawning, which many experience as contagious, is linked to such social cognitive processes (Ferrari & Coudé, 2018; Sternberg, 2016). To understand and empathize and/or mirror others have effects on our own motor system, thus bodily states would contribute to the building blocks of social aspects of our lives, imitation and language acquisition (2013, 321). Humans can use their self-experienced actions, emotions and thoughts to understand the actions, emotions and thoughts of others (Oberman & Ramachandran, 2007), and visual information of action may be mapped not only onto our visual analysis but also our motor representation in our nervous system (Rizzolatti et al. 2001). Interestingly, people on the autism spectrum, who have various degrees of social and communicative deficits (2007), and have

reduced Mirror Neuron System activity during social mirroring (Iacoponi & Dapretto, 2006), also have less body language, and have various degrees of motoric clumsiness and slowness in their psychomotor tempo (NHI, 2023).

- Bodily movements in cognitive tasks:

When performing cognitive tasks that involve remembering, problem solving or imagining something, our bodies can be used to offload information and make the task easier in the cognitive processing. This can be done by body postures or facial expressions, and by holding these casually or constitutively facilitate both access to and retention of memories.

We have “online embodiment” and “offline embodiment” to describe two different grades of bodily involvement in cognition. The former is when the body is directly involved, while in the latter, the body is indirectly involved, by way of neural simulations. Online embodiment can “evolve” into offline embodiment at a later time, by storing actual embodied responses to later be used in offline processing (2013, 321). This in turn makes bodily responses and experiences a central part of our conceptual knowledge.

Further, Foglia & Wilson refers to how the body can function as a constraint on cognition. That is, the way our bodies are constructed and interact with the world, forms the way we think about it. Examples are the ways in which our eyes and ears are designed: two of each, our eyes’ retinal cells have specific properties (while other animals have other properties), they only point forward, with some room for movement, and are essentially “inside” the head, and our ears have a certain distance between each other, and point only in one direction (unlike some animals which can move their ears, and are, say, on top of their head). We are upright animals on two legs, with two extra limbs to do other things than keep balance.

The latter features mentioned makes us think in terms of “up” and “down”, and we have a “front” and a “back”, while a long and flat creature would not conceptualize these things. And we may naturally consider that blindness and deafness changes perceptions of the world. These kinds of constraints do not necessarily contribute to mental processing, but our way of constructing and using metaphors are heavily based on our embodied experiences, drawing from George Lakoff and Mark Johnson’s book *Metaphors We Live By* (1980). Although Lakoff & Johnson focus on the metaphors of the English language, metaphors are said to be known in all

languages (Lakoff & Johnson, 1980). They argue that central cognitive processes, such as those concerning space and time, are heavily influenced by metaphors in our language.

Foglia & Wilson, with rooted examples from Lakoff & Johnson, argues that metaphors reflect the embodied experience that we have as creatures that move through (and experience and interact with) the world in particular ways. This again is reflected in our language, like embodied responses in relation to sentence comprehension, like understanding a sentence is affected by the text meaning's relation to the body's biomechanical features; if it is compatible, judgment of meaning is faster and more accurate.

Finally, they suggest that it is very possible that "body-induced changes regulate brain enhancement, information processing, and cognitive development", suggested by the increase of cortical movement representations that follows the learning of behavioural abilities (2013, 322). It would be interesting to look further into how various body-induced changes can affect, in this thesis's particular case, learning, processing and developing new language skills.

In Ionescu & Vasc (2014), they go into how EC is connected to education and better learning. As post-cognitive approaches rise, this appears as a bit of a challenge to psychology, and so to how we approach educational systems. Like with Foglia & Wilson, Ionescu & Vasc presents cognition as dependent on the body, as well as on context (see Ionescu & Vasc, 2014 references). They use the term "the morphology of our bodies" as one factor that constantly influences our cognition. By having the body and mind interconnected, cognition is as dependent on the body (as well as context) as the body is of our cognition. In the traditional study of cognitivism, the two function in different ways, separately, where cognition has been seen as symbolic processing (Bickhard, 2008; Pylyshyn, 1980). Ionescu & Vasc give reference to a multitude of studies in a wide variety of areas, including language comprehension and language learning, that show that sensory-motor processing to be an intrinsic part of higher cognition, and that acting and sensing should be considered as part of thinking itself (2014, 276), and thereby a part of the very definition of cognition.

It is shown that the initial state of the body in a learning situation have an effect on the implementation of our mental representations, in that they seem to be mental simulations of said state of the body (Barsalou, 2003; Borghi, Glenberg, & Kaschak, 2004; Glenberg et al., 2008; Ionescu & Vasc, 2014). It has been shown that "transfer sentences" involving both concrete *and*

abstract objects (e.g. concrete object: pizza, abstract object: argument) triggers higher motor provoked potential than non-transfer sentences, implying that language comprehension activates motor areas of the brain, no matter if the sentence is concrete or abstract (2008; 2014).

There is also more evidence that gesturing facilitates children's learning in equivalence tasks, like discovering better strategies when imitating "correct" gestural grouping of the equation. Looking at children's spontaneous gesturing when trying to solve tasks like equivalence tasks (and other types of tasks and concepts of a wide variety) can reveal knowledge that the children cannot yet verbally articulate to solve the problem. Studies on writing techniques also show that writing by hand has a positive effect on letter recognition in reading tests compared to using typewriting (of which I assume many teachers will be happy to hear), suggesting that sensory-motor experiences has a facilitating effect on learning (Kiefer, & Trumpp, 2012; 2014). To directly quote Ionescu & Vasc, "the representation of the concept of transfer is grounded in the motor systems of the brain" (2014, 277).

Essentially, by arguing that the body and the mind are closer intertwined, we can assume that the two are working together to process information, both received and expressed, optimally. This intertwined relationship could then be used both for easing up our cognitive load when processing new information, or in some cases, may make the processing more difficult, if the cognitive and the sensory-motor system are not working properly together. What is also clear is that it is the *appropriate* sensory-motor sensations/experiences/movements to facilitate processing, while using movements that are in some form dissociative from the cognitive processing can have little to no effect, or at worst stagnate the learning process. This in turn could have a significant effect on the way we're learning in institutions of different kinds, both where physical learning is already essential, but also in the abstract and theoretical, like mathematics or linguistics. By recognizing that we embody our cognitive knowledge, we can further look into what different ways we can use EC to facilitate learning (and even make it more fun), as well as seeing how it affects us.

Dwight Atkinson argues that *embodied* and *extended* cognition both can be seen as a single synthetic experience, as a contrast to cognitivism. He describes extended cognition as a view that the cognitive mind is inextricably tied to the external world and environment, while EC views our cognitive brain activity as something rooted in the body, in both its actions and its static states (Atkinson, 2010, 599). These essentially may work together, as the body links our mind to

the world around us, experiencing it and learning from it through our bodies. Cognition, and so our mental processing of information, depends heavily on our external environment, like doing math with calculators or on paper (or, with direct embodied inclusion, we use our fingers), our everyday activities and schedules follow clocks and calendars, arrange books alphabetically, or by colour or even size if you will, or by subject/theme. Atkinson quotes Clark (2001, 142): "The naked biological brain is just a part (albeit a crucial and special part) of a spatially and temporally extended process, involving lots of extraneural operations" (quoted in Atkinson, 2010, 601). Essentially, our brain alone cannot solve problems all on its own in a vacuum, it needs the "machine" that is our body to interact with the information it receives. Atkinson adds to the argument for EC by referring to mirror neurons that activate both in our own motor actions as well as when *others* perform the same actions (Rizzolatti and Craighero, 2004, 174; Atkinson 2010, 604). This, argues Atkinson, suggests that mirror neurons as a shared neural action understanding, for both self and other, accounts for action-oriented understanding, imitative learning and synchronized behaviour (2010, 604)

Tying EC to second language acquisition, Atkinson refers to several researchers' findings on language and our cognitive motor functions, like the brain's "language areas"" activating during sensorimotor action (Bonda et al., 1994; Atkinson 2010, 605), that our brain "motor areas" activate during speech (Hauk et al., 2004; 2010, 605) and that linguistic tasks are facilitated when accompanied by action (Rieser et al. 1994; 2010, 605). Working on a sociocognitive approach to second language acquisition, Atkinson refers to *The Inseparability Principle*, which tells us that mind, body and world work together in both learning and second language acquisition/SLA. As an example, we can see parts of a picture (a young girl, white background) and deduct some information, but not necessarily more than who we're viewing. Then, as we get more and more information in the picture (adding more of the young girl, holding a pencil and a paper lying in front of her, then seeing an older woman besides her also writing, then seeing that they are, most likely, at home etc.), slowly gives us more and more information to answer questions of what the person is doing, where they are, maybe what time it is, and maybe even answer some why-questions about the situation in the picture. By gradually learning more through visuals, Atkinson argues that learning is contextual, that the environment and contexts affect our learning, and by then could not be optional extra information (2010, 609). Also, in learning connections between signs and interpretations, we not only process like a computer, but find value and meaning in the information and learning we gain. This, says Atkinson, affects learners' engagement with learning opportunities (2010, 610).

Finally, Atkinson argues that if cognition and learning is this complex and multimodal, it must be studied in complex and multimodal ways, rather than being forced into a monomodality. Instead of detaching knowledge from the world and internalizing it, we can focus on learning as relational, both to people and our environment, and as experiential, participatory and, naturally, guided (2010, 610-611). This can be applied to, among many other things, language learning and SLA in and out of classrooms. Language learning participants and study fields could be taken out of the office and laboratories, and see what comes of the multimodal and complex way of learning "in the real world". Naturally, learning in a classroom will always be a big part of language learning, and should not be taken away, but integrating the principles of inseparability, and that our information processing can be eased by the help of different kinds of embodiment. Here, taking into consideration what embodiment can possibly be, the list can almost seem endless, so the challenge is to narrow it down to different experimental methods, focusing on relation between the cognitive and the external actions or environment.

What the study of EC shows us is that the mind and body are intertwined in many ways, and the mental activity is at times even referred to as *based* in the body. In addition, the cognitive mind is not only tied to, or based in, the body, but tied to our external, physical environment, which we are linked to through our bodies, and how we interact with it. Previous research shows that concepts like gesturing, visual perception, bodily movement, mirror neurons etc. all contribute to the theory that our bodies are highly important in cognitive processing of information. Gesturing, and other types of movement, can both facilitate learning and recognition of acquired knowledge. We understand both concrete and abstract concepts through embodiment, and our language comprehension, as well as speaking, activates motor areas of the brain, and linguistic tasks are facilitated when accompanied by action, meaning language and motor functions are tied together. This precisely will be the base theory of this thesis, by looking at how action can help language learning. Since action is embodied, and action is central to theatre and acting, this could potentially be a tool to facilitate language learning. From here, I will present how different empirical studies have used different methods that can be tied to embodiment and EC to facilitate learning language.

2.5 Empirical studies

In this section, I examine some empirical studies tied to EC and language learning. The main focus within language learning in these studies will be pronunciation, accent, accentedness, and comprehensibility. Subsection **2.5.1** will cover explicit pronunciation training, involving form-focused instruction and corrective feedback, as well as visual feedback. The studies further presented in **2.5.2** all contain some form of EC, which include: hand gestures, hand clapping, finger tapping, tactile cues, and the use of body language and prosody. Then, I will consider the role of action on L2 pronunciation training in **2.6**, which will be exemplified in greater detail in the included studies. Finally, I will present the phonological systems within Danish and Norwegian in **2.7**, presenting similarities and differences in the phonological inventory, and what the possible difficulties of learning Norwegian as L2 for a Danish L1 speaker could be.

2.5.1 L2 pronunciation training

Previous research on pronunciation reaches far and wide, with studies focusing on form (FoF), also known as form-focused instruction (FFI), and FoF plus corrective feedback (CF) and audiovisual input. Form-focused instruction (without CF) can facilitate learners' perception of new phonetic and phonological categories, and (with CF) can give some significant gains under trained lexical conditions (Saito & Wu, 2014), and can improve learners' pronunciation development in both controlled speech and spontaneous speech (Saito & Lyster, 2011). Both teachers and students see benefits of both isolated and integrated instruction, for both overcoming L1 influence on the target language, and improving fluency and automaticity (Spada & Lightbrown, 2008), and group member feedback can improve attitudes and pronunciation abilities (Lord, 2008).

Another method used as a tool is visual feedback, an enhanced method of FoF, as a pedagogical method for teaching pronunciation at the segmental level has been shown to give significant improvement on L2 production (Olson, 2014), and suggests better results for pronunciation and production than auditory learning alone (Hardison, 2003), and helps phonetic perception (Summerfield, 1979), and produced voice onset time on short controlled utterances, as well as expanding into more continuous and spontaneous speech (Offerman & Olson, 2016), and production training alone can additionally improve perception, with differing rates of improvement between individuals (Kartushina et al., 2015, this particular study also looked at Danish vowels).

In other studies focusing on other specific types of measurements of pronunciation, explicit phonetic instruction can significantly improve comprehensibility (Saito, 2011) and pronunciation on specific features in focus (Lord, 2005), and suprasegmental-based instruction can lead to gains in phonological development, like overall comprehensibility, word stress, rhythm and intonation (Saito & Saito, 2017).

Finally, different types of explicit pronunciation instruction can yield positive results, like explicit pronunciation instruction with focus on segmentals and suprasegmentals showing significant progression in pronunciation, measured by comprehensibility (Zhang & Yuan, 2020). However, when it comes to the production at the spontaneous level, as well as spontaneous speech at delayed posttests, groups focusing on suprasegmentals show significantly better results than segmentals. Interestingly, there was no particular focus from the instructor on individual vowels and consonants in the suprasegmental group, and still they had greater gains (2020, 5).

2.5.2 Embodied Techniques

In this section, I will refer to a wide body of work on different embodied techniques and their effects on language learning.

Hand gestures

There is a growing body of empirical research shedding light on embodied learning/training on L2 with the aid of gesture. Beat and metaphoric gestures both show evidence of improved learning of lexical stress, albeit the presence or absence of a written accent (in this case, Dutch vs. Spanish) could provide a greater prediction on learning lexical stress rules (van Maastricht et al., 2019). Beat gestures expressing prosodic prominence can favour information memorization in L1, novel words in learning a second language, as well as lead to pronunciation improvement in L2 (Kushch, 2018). It also gives better accentedness ratings for non-natives on more discourse heavy items (Gluhareva & Prieto, 2017). Both Kushch's and Gluhareva & Prieto's tests involved short amounts of training (15 minutes and 30 minutes respectively), but acknowledges their contrasts (and possible limitations) to other studies that are more longitudinal.

Studies on durational hand gestures show similar results. In looking at whether horizontal durational gestures can improve perception and/or production of vowel-length contrasts in

Japanese, it shows that while perception didn't improve significantly in either context, pronunciation accuracy was significantly improved by producing gestures (Li et al, 2020). Also with Japanese, although with no significant differing effect on productive knowledge, there was an effect on receptive knowledge. Hand clapping accompanying the segmentals was then shown to improve memory of pronunciations over time (Iizuka et al., 2020). There has, however, also been studies concluding that observing and producing gestures has no significant effect on learning vowel distinctions, possibly because of auditory distinctions within a syllable at the segmental level, which are very small units (Kelly et al, 2014), and while gestures might not improve perceptions, observing mouth/lip movements in training significantly improves perception of difficult L2 phonemic contrasts (Hirata & Kelly, 2010). Hirata & Kelly have observed several instances where gesture seems less beneficial than many other studies show, therefore bringing a variety into the study of gestures in aiding L2 learning. They found that co-speech gestures representing phonemes did not affect L2 learning in a significant way, despite aiming for making the gestures more "Japanese-like" and embodied through producing the gestures (Kelly & Hirata, 2017). They do, however, acknowledge that noisy data had a rather large number of participants that was excluded, and that the data was less clean than typical auditory FL paradigm.

Finally, there are still other studies that favour gesture, for example in relation to pitch. Gestures conveying pitch contours in Mandarin can strengthen learners' representations of Mandarin lexical tones and word meanings (Morett & Chang, 2015), both in observing and producing in facilitating tone identification and word learning (Baills et al., 2019) While semantic gestures hinder tone identification (2015), observing and producing pitch gestures can trigger enhancement of both tone identification and word-learning results (2019).

Hand clapping and finger tapping

In studying hand clapping, there is evidence that suggests that hand clapping can enhance pronunciation skills in a foreign language, like the duration of word-final vowels, by clapping to the rhythmic structure of cognate words in French and Catalan in 7-8 year old children (Baills & Prieto, 2021). Lee et al. (2020) looked at pronunciation instructions using either perception- or production-based form-focused instruction (FFI), and found that clapping, as well as describing tongue and lip formation while demonstrating pronunciation, had the greatest effect on pronunciation gains. Finally, hand clapping with Chinese learners of French seemingly has given

only a near-significant effect on accentedness, though acoustic durational measures had great improvement (Zhang et al. 2018).

In Baills & Prieto (2021), the young learners focused exclusively on pronunciation, rather than including word meaning, to hinder cognitive overload. It is possible, says the authors, that this pronunciation only focus, along with using only cognates where meaning was more easily acquired, is the reason the results were significantly better for the clapping group, while in Zhang et al. (2018), and also Izuka et al. (2020) found less effects on hand clapping, but in these studies, students had to learn both pronunciation and word meaning at the same time, possibly giving less focus and effect to the pronunciation. Zhang et al. (2018) did find hand clapping to get higher scores than non-hand clapping as well as durational realization of the last stressed syllable of the word. They hypothesised here as well that cognitive overload from learning both new word meaning and pronunciation. Izuka et al. (2020) also involved both vowel and consonant duration patterns, with Japanese students learning English, increasing the difficulty and cognitive load, as opposed to only vowel duration, and two languages that share more traits, like French and Catalan.

For the study from Lee et al. (2020), of the four groups of learners, the groups featuring syllabic perception instruction group (SPe) and phonetic perception instruction group (PPe), had the greatest gains across four groups altogether, including past the post tests. The SPe group practiced rhythm, stress and syllables using clapping, and the PPe group used descriptions of tongue and lip formations and demonstrated minimal pairs. Although students in the SPe still had some trouble identifying what constituted as a syllable, especially with longer items, their success rate of picking up suprasegmental stress patterns of the items was quite high.

Tactile cues

Finally, the use of tactile cues and haptic pronunciation teaching (the use of movement and touch) shows promising results, like with “The Butterfly Technique”, with tapping of the right shoulder when uttering a stressed syllable, and tapping the opposed forearm/elbow for an unstressed syllable (Burri & Baker, 2016) or other multiple techniques that encompass a combination of tactile and kinesthetic learning to teach several types of pronunciation, as opposed to clapping/tapping that focuses mostly on stress and rhythm (2016; Acton et al., 2013; Teaman & Acton 2013). In addition, these types of tactile learning techniques, learners report

them being more fun and engaging, increasing enthusiasm to learn more (Burri & Baker, 2016, 2019; Acton et al., 2013). Tactile cues accompanying pronunciation training involving Turkish learners of English practicing /θ/ and /ð/ while touching their own tongue when protruding from the lips, gives a greater improvement than non-tactile cues on discourse reading taks (Ozakin et al., 2022). The word- and sentence-imitation tasks had no significant positive improvement with tactile information. Ozakin et al. (2022) empirically tested the effects of tactile cues, and references Esteve-Gibert et al. (2019), reporting that adding tactile information had no significant effect on L2 perception. This was on 5YO Catalan/Spanish bilinguals, learning the English /æ/–/ʌ/ vowel contrast, watching themselves in a mirror, touching the lips to feel the rounding difference. Ozakin argues, however, that this method could be unfavourable, because of /æ/–/ʌ/ also having difference in tongue position, and that the learners had no visual input from a native speaker.

The task difficulty could be a possible account for the difference, in that the imitation tasks included a native model speech, while relatively low scores in pretest could give room for improvement for the discourse tasks (2022). By referencing Llompart & Reinisch's (2019) work, Ozakin et al. discusses imitation as being not as demanding as discourse reading, and that though imitation abilities are related to learners' phonological representation of L2 sound contrasts, they don't necessarily reflect learners' productive skills in using the same L2 sounds (2022, 2019). Imitation often does entail a certain aid in repeating a native sound (either in language, or in a way of speaking, or a type of voice etc.), while different kinds of reading tasks leaves the learner more up to themselves, where a stronger bodily/embodied memory, like learning exactly where to place the tongue on certain phonemes, could benefit further improvement of pronunciation learning.

These studies suggest that a form of embodiment during language learning has a positive effect on pronunciation training like accentedness, particularly stress, intonation and duration of vowels when using clapping or tapping movements. Some of these embodied movements could be limited to pronunciation training that can only more or less directly depict the items to learn, while others can go beyond the segmental level, and broaden the types of pronunciation training needed for learners (see a full, quite interesting and inspiring, list of different protocols by Teaman & Acton, 2013 in the references list). In the next section, I introduce and talk about a study by Bach-Marqués and Carrera-Sabaté involving prosody and gestures.

Body language, prosody, and gesture

Bach-Marquès and Carrera-Sabaté (2019) conducted a study involving learners of Catalan, with the help of prosody and gestures. Learners watched videos of native speakers, speaking out short phrases, with both gesture and prosody involved. They theorized that body language and gestures could improve the production of phonetic elements.

Results showed the participants reporting during feedback that the sense of different body positioning and movement during the utterances helped enhance their pronunciation (2019). Gesture use in general, either observed, performed, imitated, resulted in a significant raise in accuracy of appropriate productions, going from 45% without gestures and body movement, to 85% when including gestures. The results showed particularly high accuracy when the gestures resembled the drawing of the intonation contour. The researchers concluded that L2 sounds would be easier to pronounce with aid by body movement and proprioception, by making pronunciation visible and tangible.

This study is a contribution to the notion that gesture, physical action and EC might contribute to better pronunciation. This is still with gestural instruction, but including the prosody, and some of the videos involved a certain kind of "involvement", enhanced by the prosody. The combination of body language, prosody, and this "involvement" in speech is what I refer to as "action" in this thesis. The components within can be actions themselves, though they rarely appear alone during speech, making the use of action in language learning research one way to "mimic natural speech", so to speak. In the next section I will get further into the theme and use of action during L2 pronunciation learning.

2.6 The role of action on L2 pronunciation

Here I will present some studies that focus on the role of action on L2 learning and pronunciation. I will first present two studies that used drama in a language learning setting (Abenoja & DeCoursey, 2019; Galante & Thomson, 2017) and how it affected the pronunciation of the target language. Then, I will present two studies that focused on the use of surprise to help facilitate and improve "the audio-phonatory behaviour" (Estrada, 2004, 2007).

Abenoja and DeCoursey (2019) examined Chinese students enrolled in a beginner's French elective course at a Hong Kong tertiary institution in 2014, where drama was used as an approach to teaching French (conducted mainly in English and occasionally in French). They found that using these activities made the target language easier to recall and learning more enjoyable. Lessons involved physical warm ups to make students active, conducted purely in French, changing classroom layout from traditional sitting classrooms, to a standing classroom (with tables and chairs moved out of the way), pronunciation of the French alphabet, and role-plays and dialogues for students to memorize and dramatically play out. Many students reported this as relaxing, useful, interactive and fun, as well as making homework easier than other types of homework. Abenoja and DeCoursey wrote that from the students' reports and reflections written in diaries after classes, their responses gave weight to the theory that drama activities strengthens and helps learning in the classroom, like students remembering better what they learned, possibly because emotional ties to the phrases they use during lessons (p. 727).

The study is like many others in that it was conducted with students from a traditional study setting, where medium of instruction (MOI) involves mostly sitting down, practicing grammar and pronunciation as the common method for language learning. In this case, it is also within Confucian-heritage cultures, or CHC, a term popularized by Carless (2011) and Biggs (1996), meaning that the drama setting was a definite differing method for the students in the class.

Drama as a method in education has a longer tradition than one would assume. Abenoja and DeCoursey refers to Dorothy Heathcote and Gavin Bolton's work in the 1970's and their focus on "learning through doing". Heathcote pioneered "process drama", which can be defined as a creative pedagogical practice that approaches language teaching and learning through drama, where methods like scripted performance, 'role play', meaningful context for spontaneous language production, and co-constructed, imagined drama worlds are used (Hulse and Owens, 2019).

In this particular study, the researchers wanted to look closer at the effects on L3, as opposed to the more commonly studied L2. They asked how drama activities benefited the students' language learning of L3 in the classroom, and how they perceived the drama activities as a language learning approach. The results, as mentioned, were mostly based on the students' own

reports and opinions of the layout of the class. Reports were that it was “more fun”, that they were allowed to “speak more”, that the warm ups “woke up the brain”, and there seemed to be a greater enthusiasm and willingness to actively use the language. One student reported that playing a role plays a positive effect on her oral skills, by helping her “speak the words correctly and with emotion”. Most of the reports from the students were about enjoyment and remembering the vocabulary and grammar, while not many reports from anyone involved were about the effect on accentedness, although through the interviews and reflective journals there was indications that “correct pronunciation of French words was also easier to remember” (p. 729). The students reported that the drama activities gave them the opportunity to “check their pronunciation” with other peers, as opposed to practicing alone. The latter results does suggest that there are some effects on pronunciation, but accentedness and comprehensibility are not mentioned, meaning we don’t know exactly how the pronunciation were affected, if they managed to improve their accentedness, or if comprehensibility or intelligibility were positively affected.

A recent study by Galante & Thomson (2017) looked into how instructional language teaching techniques adapted from drama could positively affect comprehensibility and accentedness in Brazilian learners of English, as well as fluency. It involved a pretest-posttest design, with one group receiving drama-based English instructions and one control group. The results showed that the drama group had significantly higher gains in L2 oral fluency than the traditional EFL group. However, comprehensibility scores were only slightly affected, and differences in accentedness did not change in either group (2017, 115).

Unlike many pronunciation studies which use reading tasks, Galante & Thomson employed tasks with mostly free-form narration, in first person, third person and a video narration, in addition to role-play and a monologue. They also ended up performing pre-written 15 minute plays for a familiar audience at the end of the 4 month program. For specific pronunciation training, the L2 learning included supplementary pronunciation practice for the Brazilian learners of English, which included practicing and listening on both segmental and suprasegmental features. In rehearsing the short plays that the students were to perform, character development, and understanding the meaning of the lines and how to convey that meaning was the main focus, not rote memorization. This in turn gave room for improvisation and changing of the lines.

The comprehensibility scores were affected to some degree for the learner group, but not significantly better than the control group. The researchers discuss the possibility that the improvement did not come from their laboratory classes where explicit pronunciation was the focus, but possibly rather because of paralinguistic features of speech that were incorporated during the drama classes, like expressing emotions and practicing vocal projection and volume (2017, 133). As the researchers say, a single variable cannot explain such a small difference in comprehensibility improvement, but the drama-based approach may have facilitated the practice of pronunciation and its uptake by the learners. As mentioned, when it came to the accentedness rating, both groups received lower accentedness ratings over time, but the drama group had no significant advantage. The raters were Canadian English speakers, which the students did not receive specific instruction in, which could have affected the ratings somewhat (2017, 133). Further, although the learners' speech was characterized as strongly accented, they were rated as comprehensible, adding to research suggesting that accent and comprehensibility are somewhat independent of each other, and that heavy accent does not automatically entail low comprehensibility (2017, 133-134).

It seems that in the specific context of drama-based lessons had a positive effect on pronunciation training, more so on comprehensibility than accentedness. Although no findings on significant accentedness improvement was found, theatre and drama could still contribute as a tool within accentedness improvement. First person narrations, placing the learner in an acting role, shows signs of giving less accentedness ratings and better comprehensibility ratings (2017, 134). The dialogue role-play from this study could be interpreted as less personal, since it was about one person welcoming another to their country, possibly leading to more of a descriptive dialogue than a personal one. Suggestions of first person focused text improving accent and comprehensibility shows possibilities of further exploration within the use of theatre, acting and drama as pronunciation improvement tools.

Research form Estrada (2004, 2007)[1] looked at the effect of *surprise* to facilitate the learning of French front vowels for Spanish and Catalan speakers, using a "verbo-tonal approach" (French: *L'approche verbo-tonale*), with consideration of extra-linguistic factors, in this case, emotions and the melodies and tonalities that follow (2004, 319). The basis of the research leaned on the

[1] As a warning to the reader, this section is based on two papers written in a somewhat advanced French. I, the researcher, have somewhat limited knowledge of French, but have tried my best to translate, checking online and physical translators/dictionaries. If some translations seem off, or not precise enough, criticism is taken

proposition that speaking is a dual "audio-phonatory" activity, never separated from hearing, and that the perception of sounds in speaking situations has specific perceptions of the sounds, rather than being general (2004, 320). One goal of the verbo-tonal approach Estrada uses is to establish the optimal situations and contexts which will be facilitating to spoken communication, and in addition, for integrating a new phonological system for learners, where contexts could include different "values" of spoken language (French: *les valeurs de la langue parlée*), which includes tempo, rhythm, melody and what they call *le timbre des sons,* or the timbre of sounds (2004, 320).

Comparing surprise to two other speech acts of "demanding information" (French: *demander d'information*) and "giving an explanation" (French: *donner une explication*), with no difference in the utterances themselves, measuring the melodic relief, the duration of the statements, the tonal dynamics and the last two tonal or melodic variations of the statements (2004, 322), they found that the manifestation of surprise was characterized by a lengthening of the utterance, by a single accent in the predictive part of the utterances, and by a final melodic variation located in the over-acute register (2004, 326-327). While interrogative utterances proved to be the shortest in duration than the other two speech acts, the surprised utterances proved to be the longest. The use of surprise also gave a greater tonal dynamic than the other two speech acts considered (2004, 324). The use of surprise also constitutes a positive effect on the production of high front vowels and central vowels (2007, 91). The results suggest that the use of surprise makes up for one of the most acute realizations of speech, and could be optimal in conditioning speakers for more precise speech of L2 (2004, 327). Estrada elaborates on specific challenges for the participants in the study, in that while Spanish-Catalan speakers have the tendency to "ultrasegment" or "over-emphasize" utterances in French, which could have effects on the realization of phonemes, the expression of surprise would be likely to contribute to a neutralization of such a tendency (2004, 324; 2007, 89).

The study shows that the use of expressed emotion aids and facilitates certain prosodic and phonematic behaviour, and makes speech productions more precise, constituting optimal conditions for the structuring of foreign phonemes (2004, 327). Using a focus tool such as emotion, going beyond mere the use of perception and production practice and the strictly linguistic framework, can indeed be a useful pedagogical tool. Estrada also points out that the use of *surprise* was chosen as opposed to *astonishment*, two emotions very divergent in terms of the illocutionary force conveyed, in the sense that surprise is of communicative nature and of

dialogical character. By manifesting the emotion of surprise, an appeal to the interlocutor is made, who also will seek to obtain information concerning the triggering element of the emotion (Estrada, 2007, 87). Astonishment, on the other hand, can be a monologue-oriented linguistic manifestation in the absence of an interlocutor. Estrada does not reject astonishment as an affective manifestation to be used other than it not fitting into the study's conversational focus.

Like Estrada points at, the differing audio-phonatory behaviours between speakers (and intra-speaker) in relation to the emotional speech of surprise, is to be taken into consideration in a pedagogical context, in that there needs to be a diagnosis of the speaker's achievements in order to optimize the pedagogical approach (2007, 91). Individual differences are always to be expected, and to be approached with some level of consideration for the individual behaviour. The use of a prosodic-affective realization such as surprise indeed can have a positive and facilitating impact on both final phonetic realizations and the structuring of the whole vowel system for a specific learner profile (2007, 91). By using emotion, the learner can employ an embodied engagement in speech, which then facilitates their learning and pronunciation of L2.

Studies clearly show a positive effect of drama and the use of emotion-focused speech. In classroom situations, the use of drama demonstrates that students find the education more fun and engaging, while feeling more confident in their speech production. Pronunciation is affected as well, and comprehensibility scores show an increase. Accent shows improvement through the use of drama in the classroom, though not necessarily significantly more than the control groups. However, the use of emotions like surprise in dialogue shows a significant improvement and generates a facilitating effect on pronunciation. What is not mentioned in the latter case is how the accentedness overall improved long term, only the segmental phonemes of learners of French. The improvement of both engagement in speech and the segmentals could possibly give overall a better accentedness and comprehensibility score. This will be of focus in the study I will describe in Chapter 3 on the methodology. Based on these studies, the triggering of emotions, and focusing on speech from a first person focus, might improve pronunciation overall, as well as improve the accentedness through improvement of the realization of phonemes.

2.7 Danish and Norwegian - Sister (or rather cousin) languages

Between Danish and Norwegian, there are obvious similarities, with a close to identical writing system (Standard Bokmål and standard Danish writing). The languages are mutually intelligible, though with some asymmetry, where Norwegians more often understand Danish than the opposite (Delsin & Åkesson, 2005). A few reasons for this could be more exposure to the Danish language to Norwegians than the opposite, especially the older generations, with Norway having a tradition for using written Danish for hundreds of years, and assimilating to the language, leading to large similarities in vocabulary (Torp, 2006). Linguistically, Danish is closer to Swedish as East Nordic languages, Norwegian is closer to Icelandic and Faroese (and the extinct languages Norn and Greenlandic Norse) as West Nordic languages (Basbøll, 2005, 6). However, as mentioned above, history has brought the two languages closer in mutual comprehensibility.

I will present the relevant similarities and differences in this section. In 2.5.1 I will go through the phonological inventory of both languages, with Danish vowels, Danish consonants, Norwegian vowels and Norwegian consonants, in that order. In 2.5.2. I put the two different phonological inventories in a grid to highlight the similarities, and most importantly, the differences, to establish what phonemes a Danish learner likely could struggle with.

2.7.1 The phonological systems of the two languages

Danish has a particularly large vowel inventory compared to Norwegian. It is very common for unstressed syllables to be considerably reduced in informal rapid speech. There is also an unusual prosody which gives little information about sentence structure (Grønnum, 2008).

Hans Basbøll sites these as Danish's vowel inventory, totalling of 27 vowels:

> Short: i, e, ɛ, a, ɑ, y, ø, œ, ɶ, u, o, ɔ, ʌ, ɒ
> Long: iː, eː, ɛː, æː, ɑː, yː, øː, œː, ɶː, uː, oː, ɔː, ɒː
> (Basbøll, 2005)

It should also be noted that schwa /ə/ is a common feature in Danish, as a result of consonant and vowel reduction (for more on schwa, see Basbøll, 2005 and Haberland, 1994). There is no mention of the use of the use of /ɪ/ and /ʏ/ for short vowels.

Table 1

Danish Vowels

	Front	Mid-front	Mid	Mid-back	Back
	unrounded/rounded				
High	i / y				u
				ʊ	
Mid-high	e / ø				o
			ə		
Mid-low	ɛ / œ				ʌ / ɔ
	æ		ɐ		
Low		a / œ			ɑ / ɒ

Note. Adapted from *The Phonology of Danish,* by Hans Basbøll, 2005, New York: Oxford University Press

The Danish consonant inventory is somewhat simpler. Hans Basbøll includes 18 consonants, on which I base Table 2:

Table 2

Danish Consonants

	Labial	Alveolar	Palatal	Velar	Uvular/ pharyngeal
		Unvoiced / Voiced			
Stop	p b	t d		k g	
Nasals	m	n		ŋ	
Fricatives	f	s	ɕ		h ʁ
Approximants	v	ð	j		
Lateral		l			

Note. Adapted from *The Phonology of Danish,* by Hans Basbøll, 2005, New York: Oxford University Press

NORWEGIAN

Norwegian has a considerably smaller inventory of vowels than Danish. Gjert Kristoffersen in his book The Phonology of Norwegian (2000) sites these vowels:

Table 3

Norwegian Vowels

	Front	Mid	Back
	Unrounded / Rounded		
High	i / y	ʉ	u
	(ɪ / ʏ)		
Mid-high	e / ø		o
		ə	
Mid-low	ɛ / œ		ɔ
	æ		
Low			ɑ

Note. Adapted from *The Phonology of Norwegian*, by Gjert Kristoffersen, 2000, New York: Oxford University Press

Although he uses /i/ and /y/, [ɪ] and [ʏ] can be seen in Norwegian transcripts as well. Kristoffersen does also claim /æ/ to be a marginal phoneme, saying it patterns as an allophone of /e/. I would argue that many native Norwegians would absolutely claim /æ/ to be a phoneme, since it is an integral part of most, if not all dialects of Norway. As an example, John Ole Askedal, in *The Germanic Languages* in the chapter on Norwegian, includes /æ/ as a definite part of Norwegian phoneme inventory, despite also mentioning the possibility of it being an allophone of /e/ (Askedal, 1994, 221). If we then include /æ/, he thus cites 9 short vowels and 9 long vowels, totalling 18 (Kristoffersen, 13, 2000).

Further, Norwegian has 24 consonant phonemes, presented in Table 4. I have adapted the grid from Kristoffersen, with some extra information added:

Table 4

Norwegian Consonants

	Bilabial/ labiodental		Dental/ alveolar		Retroflex		Palatal	Velar		Laryngeal
	Unvoiced / Voiced									
Stop	p	b	t	d	ʈ	ɖ		k	g	
Nasals		m		n		ɳ			ŋ	
Fricatives		f		s		ʂ	ç			h
Tap				ɾ		ɽ				
Approx.	ʋ	w		l		ɭ	j			

Note. Adapted from *The Phonology of Norwegian*, by Gjert Kristoffersen, 2000, New York: Oxford University Press

2.7.2 Differences in phonological systems

Now that we have our consonants and vowels, we can compare the two in a gathered grid system in Table 5. Here, regular black text is the shared inventory, while green is Norwegian only and Red is Danish only.

Table 5

Norwegian and Danish Consonants Compared

	Bilabial/ labiodental	Dental/ alveolar	Retroflex	Palatal	Velar/ uvular	Laryngeal
	Unvoiced / Voiced					
Stop	p b	t d	ʈ ɖ		k g	
Nasals	m	n	ɳ		ŋ	
Fricatives	f	ð s	ʂ	ɕ ç	ʁ	h
Tap		ɾ	ɽ			
Approx.	ʊ w	ʋ l	ɭ	j		

Table 6

Norwegian and Danish Vowels Compared

	Front	Mid	Back
	Unrounded / Rounded		
High	i / y	ʉ (ɪ / ʏ)	u ʊ
Mid-high	e / ø	ə	o
Mid-low	ɛ / œ æ	ɐ	ʌ / ɔ
Low		a / œ	ɑ / ɒ

We can assume that both speakers would have no problem pronouncing the shared features. Based on Flege's research, we could expect the existing sounds in Danish to somewhat interfere with, or even block, the sounds in Norwegian that do not exist in Danish, possibly creating some approximate sound category. Norwegians learning Danish would be expected to struggle with the consonants /ð, ɛ, ʁ, v/, and the vowels /ʊ, ʌ, ɐ, a, œ, ɒ/. While Danish people learning Norwegian would be expected to struggle with the consonants /ʈ, ɖ, ɳ, ʂ, ç, ɾ, ɽ, ʊ, w, ɭ/, and the vowels /ɪ, ʏ, ʉ/. Referring to Kristoffersen and Askedal, there is some debate whether [æ] is a phoneme or an allophone of [e]. I have opted to treat [æ] as a phoneme, which means that this should not be a challenge for Danish natives learning Norwegian.

There is a total absence of retroflexes from Danish altogether, as opposed to Norwegian, which has six retroflexes. This should prove to be challenging for a Danish learner of Norwegian. There is also a clear distinction in how the two languages pronounce their R's. Norwegian has several dialectal varieties in the R, with taps, trills, and velar forms, while in Danish, the [r] is realized as a voiced uvular fricative [ʁ] syllable initially, and as a vowel [ɐ] syllable-finally (Basbøll, 2005, 64), giving Danish a more limited realization of R. Considering the similarities in Danish and Norwegian, both politically and linguistically (Kristoffersen, 2000; Basbøll, 2005), one could expect Danish learners of Norwegian to learn the language quite fast, and that comprehensibility would be high. This does not automatically entail low accentedness, as shown by studies showing that heavy accent does not mean low comprehensibility, and that listeners tend to assign stricter scores to accent (Munro & Derwing, 1995, 2020). This could mean that accentedness could be rated as lower than expected.

In the next chapter, I will cover the methodology of involving the participants of the study, as well as native Norwegian raters. I will go into details in the first paragraph below.

3 Methodology

The focus of this study has been with three native Danish speakers learning Norwegian, at a Norwegian art and drama school. In section **3.1** I will briefly talk about how the project was started and how the participants were recruited. I will then present my research design, including relevant information about the participants. In section **3.2** I will present the learning methods of

the students in school. Section **3.3** will focus on the data collection; what the data was, how it was collected, and how it was analysed in Praat. Section **3.4** will address some of the challenges on the way and discuss the limitations of the research design.

3.1. Research design

This study was conducted as a mixed-methods quasi-experimental case study of three native Danish speakers learning Norwegian. The project started as a collaboration between the art institute where the students went and UiO, as an initiative to focus on the language learning of new students from outside of Norway. I was put in contact with two of the teachers running the new language programme, and then allowed to meet the students, do some unofficial observation (no data from these classes are included here because there was no written consent at this point), before later on giving consent forms for all the students, which they willingly signed (see appendix A for consent form used). From then on, I was present in some of their language classes, and started with data collection through recordings and interviews, which will be described further below.

THE PARTICIPANTS

The three Danish speakers were young students, in their early and mid 20s, at a Norwegian art institution, and were in the same second year class, which otherwise consisted of native Norwegians. All participants knew English, and one had studied and lived in Sweden, and knew Swedish at an intermediate level.

All the participants made handwritten consent to join the study, and were informed that their identity would be anonymised. During the data collection, I have respectively called them by the anonymised pseudonyms Thyra, Margret, and Alexandrine, (the names are based on Danish queens throughout the ages).

3.2. Learning Norwegian

The students in question were all taking Norwegian classes, only consisting of themselves and one teacher, to take a deep dive into the Norwegian language, mostly focusing on the phonological system and phonetic training. Each class lasted about 60 minutes, and were done consistently over the semester, usually once a week. In addition to this, the students had class

with the rest of the Norwegian students (learning a different dialect than their native one, while the Danes continued learning Urban East Norwegian), where they utilised a wide variety of practices, both physically inclined and orally focused. Some of these practices included mimicking a recording of a person with their target dialect/language, “lip syncing” to the recording, stretching elastic bands to the prosody (this technique is also mentioned in Gilbert, 2008, see sources for a further description of this particular technique) and vowel lengths, focusing on giving meaning to sentences spoken etc.

The former type of class consisted of sitting down and segmenting singular phonemes/sounds to the students. Selected Norwegian phonemes that do not exist in Danish were particularly chosen to work on, both isolated, and in longer tongue twister-like sentences. This included several of the retroflexes in Norwegian, the close front vowels, and tonemes.

For the latter class, the goal was to appropriate the new dialect/way of speaking to a level where the student could feel a closeness to the dialect and at the same time feel free to speak, act and move naturally, organically turning into a character. Avoidance was taken not to focus too much on how to “speak correctly” in the theatre, or take away the speaker’s own original speech, but to use the language in an active way towards transformation when desired. The longer process included “stages” of listening, segmentation, imitation and appropriation (Wennersten, 2022)[2]. This class resulted in a small performance where all the students played out segments from plays in their new speech form, which I attended (and enjoyed).

3.3. Data collection

In this section I will give information on how the data was collected. First I will summarize all the data types that are a part of this thesis. After that, I will go into detail how recordings with the L2 learners were done, including the content of the instructed sentences, the setup, and a brief mention of the technical gear. This will also include a brief description of the L1 speaker recording sessions. Finally, I will describe the final rating session, done by native Norwegians, of the recordings.

[2] Documents and detailed information on these language classes were received from Wennersten by email in 2022. For further information on these classes, contact the researcher of this thesis

3.3.1 Overview of data collection

The data collection in this project consisted of four types of data. First, I made recordings of the students individually, reading out loud sentences they were given, in three sessions over four months. Second, I included one interview per student per session that focused on personal experiences and feelings around language classes and learning Norwegian, as well as filling out a short questionnaire. These were all done at the same session. Third, I collected recordings of native Norwegians pronouncing the same sentences in the same ways that the students did, for comparison. Fourth, I had three other native Norwegians do a likert scale test where they listened to the gathered recordings in randomised order and rated the accentedness and comprehensibility of each individual recording. I also made sure that the three Norwegian listeners did not know the three Norwegian speakers from the recordings, thereby eliminating any potential bias based on familiarity. I will describe in further detail below.

3.3.2 Data collection of Danish learners of Norwegian

Each recording session with each student consisted of them reading out 16 sentences constructed by me. The sentences were to be pronounced only once, unless there were some stumbling or hesitation, and with a bit of space between each sentence. The sentences were constructed based on what phonemes they were practicing in class, which was the following (say them out loud if you like):

/y/, /ʉ/, /ɽ/, /ʈ/, /ɖ/, final R

In class, this was often the focus of practice for the students, and according to their teacher was the phonemes they struggled the most with. This claim agrees with the inventory of the languages, insofar that Danish lacks these phonemes. However, Danish does have a [y], which then should not be a challenge to them. It is possible then that the Norwegian and the Danish [y] are not realized identically. I Norwegian [y] is, at least for some speakers like speakers of Urban East Norwegian (UEN), is articulated with more protruded lips than [ʉ] (Kristoffersen, 2000, 16), and that the latter is more retracted with respect to tongue body position than [y] (2000, 14). This could imply that the Norwegian realization of [y] is somewhat different from Danish, even though a direct comparison between Norwegian and Danish is not mentioned by Kristoffersen, nor by Basbøll. There could still be some challenges for Danish learners of Norwegian with this phoneme, considering that /y/ and /ʉ/ are very close in proximity, while in Danish the closest

rounded vowel to /y/ is /ʊ/, possibly creating some difficulties in separating the former mentioned vowels.

These phonemes were spread around in the 16 different sentences, where each target phoneme came up between 5-7 times in total, with the exception of final R, which came up considerably more. The final R may at times be omitted in native speech, especially in rapid speech, or if appearing in certain surroundings, like appearing before a consonant-initial word. Because of this, it is useful to have more cases of final R to examine if the absence of final R in native speech, as well as where it should be present, would be mirrored by the L2 speakers. When recording, the students were told to read the sentences in two different conditions, one "with action" and one "without action". The students were at times told to do the action-sentences first, then the non-action-sentences second, sometimes the other way around (this to try to minimize patterns and give less clues to the students what exactly they were doing). The "with action/acted" condition included the exact same sentences as the "without action/neutral" condition, but where each written sentence had written in italics a context/personal feeling beneath the sentence. Examples were "You're sick of this situation", "You think this is funny", "You want to say something heartfelt to your boyfriend/girlfriend", and "You're making fun of someone". These conditions were meant to provoke some sort of acting will in the students, an inner feeling.

Table 7

Sentences from recording sessions

Norwegian Sentences	English Translation	Target phonemes in IPA, target sounds highlighted
Les opp disse setningene, med konteksten/underteksten som står under hver setning	Read out loud these sentences with the context/underlying meaning that is written beneath each sentence	-
Jeg klarer ikke dette lenger *Du er lei situasjonen*	I can't do this anymore *You're sick of the situation*	/ɽ/, /ɾ/ k**l**arer, lenge**r**
Jeg så at du kysset henne *Du er sint på kjæresten din*	I saw that you kissed her *You're angry at your boyfriend/girlfriend*	/ʉ/, /y/ d**u**, k**y**sset

Jeg vet at han eier et kart over hele byen *Du har en lur idé*	I know that he owns a map of the whole city *You got a clever idea*	/y/, /ʈ/,/ɾ/ eie**r**, kar**t**, ove**r**, b**y**en
Han burde gå først, han kjenner skogen godt *Du er redd for skogen*	He should go first, he knows the forest well *You're scared of the forest*	/ʈ/, /ɖ/, /ɾ/ bur**d**e, førs**t**, kjenne**r**
Jeg vil at du skal gå, fordi du skremmer meg *Noen gjør deg nervøs/redd/utrygg*	I want you to leave, because you're scaring me *Somebody makes you nervous/scared/feel unsafe*	/ʉ/, /ɖ/, /ɾ/ d**u**, for**d**i, skremme**r**
Sola skinner, som øynene dine *Du vil si noe vakkert til kjæresten din*	The sun shines, like your eyes *You want to say something beautiful to your boyfriend/girlfriend*	/y/, /ʈ/, /ɾ/ so**l**a, skinne**r**, ø**y**nene
Du høres ut som et dyr *Du synes dette er morsomt*	You sound like an animal *You think this is funny*	/y/, /ʉ/, /ɾ/ d**u**, **u**t, d**yr**
Du tar alltid hennes parti *Du er frustrert på bestevennen din*	You always take her side *You're frustrated at your best friend*	/ʉ/, /ɾ/, /ʈ/ d**u**, ta**r**, par**t**i
Han ville så gjerne takle situasjonen, men klarte det ikke *Du er oppgitt over noe*	He really wanted to handle the situation, but he couldn't do it *You're upset about something*	/ʉ/, /ʈ/, /ʈ/ tak**l**e, sit**u**asjonen, k**l**ar**t**e
Hun hadde de styggeste gardinene jeg hadde sett *Du gjør narr av noen*	She had the ugliest curtains I had ever seen *You're making fun of somebody*	/y/, /ʉ/, /ɖ/ h**u**n, st**y**ggeste, gar**d**inene
Fordelene er mye sterkere enn svakhetene *Du er ivrig og glad*	The advantages are much greater than the weaknesses *You're eager and happy*	/y/, /ɖ/, /ɾ/ for**d**elene, e**r**, m**y**e
Han ser ikke sin egen verdi *Du synes dette er trist og dumt*	He doesn't see his own worth *You think this is sad and regretful*	/ɖ/, /ɾ/ se**r**, ver**d**i
Vi har ikke lenge igjen, klokka tikker *Du har fryktelig dårlig tid*	We don't have much time left, the clock is ticking *You're really running out of time*	/ʈ/, /ɾ/ har, k**l**okka, tikke**r**
Jeg hater det ekle gliset ditt *Du sier endelig det du virkelig mener*	I hate that nasty grin of yours *You finally express your true opinion*	/ʈ/, /ɾ/ hate**r**, ek**l**e, g**l**iset

Jeg så en fyr i køen, og han ligner sykt på broren din *Dette er veldig mistenksomt*	I saw a guy in the line, and he looks so much like your brother *This is very suspicious*	/y/, /ɾ/ f**yr**, s**y**kt
Jeg hører kun på folk som behandler meg med respekt *Du er avvisende*	I only listen to people who treats me with respect *You're dismissive*	/ʉ/, /ɽ/, /ɾ/ høre**r**, k**u**n, fo**l**k, behandle**r**

The students were told to not hold back in how they would say this if they were using it on stage, meaning that even if they felt it was overdone or "over acted", that was acceptable. There were no other instructions given and no form of correction during the recordings, except I told them to repeat a sentence if I noticed some stumbling, hesitation or stuttering. The recording was done in a small room, with me and the student sitting opposite each other with a table between us. I put down all the sheets with the sentences before the recording started, so that there would not be any sounds of rattling and fumbling of paper during recording. The recorder I used was a small H2next Zoom Handy Recorder, done on XY configuration. The recordings were transferred to a safe separate portable SSD.

Following the recording session, I conducted an interview, which included both the recording of their thoughts, and a questionnaire with answers to a 5-point Likert scale. The open questions were about their personal feelings towards learning Norwegian, difficulties, as well as descriptions about the language classes and how they experienced these classes. The questionnaire consisted of five questions, each with five different possible answers. Instead of gradable numbers, the answers consisted of text answers, ranging from an "agreeing" end (corresponding to a 5) to a "disagreeing" end (corresponding to a 1). The questions were all about language learning as well as using physical and mental actions, trying to keep the vocabulary understandable and directed towards the acting aspect of their learning.
An example of a question (originally in Norwegian) follows here. The top answer correspond to 5, the bottom answer correspond to 1, and this was consistent through all the questions:

To what degree do you feel it helps you improve your pronunciation to interact with your teacher/your fellow students?

- ☐ It helps a lot
- ☐ It helps sometimes
- ☐ Not Sure
- ☐ It does not help a lot
- ☐ It does not help at all

The students were told to read it through and circle the answer they felt fit the most. See appendix B for the full questionnaire.

3.3.3 Data collection of native Norwegians

In addition to the recordings of the Danish students, three native Norwegians were recorded in the same manner, meaning they were recorded doing both action and non-action, reading the exact same sentences as the Danes. Only one session was done per Norwegian. All of the Norwegians spoke the same dialect (Urban East Norwegian), with small natural variations. It was decided that the inclusion of native Norwegians would provide a baseline condition for comparison, and the L2 speech could more easily be compared to native speech to identify the errors and correct pronunciations.

3.3.4 Data analysis

When the recordings had been done, the sentence recordings were put into Praat (Boersma & Weenink, 2001). I marked “correct” and “not correct” for each target phoneme in Praat and saved the annotation into a TextGrid file. Then I used a script (Elvira-García W., 2018) to export the annotated texts, which would later be used as the data points to calculate accuracy rate. The same procedure was done for both non-native and native data.

After recording sessions, the recordings were analyzed in Praat, where annotations to the target phonemes were made (here for repetition: /y/, /ʉ/, /ɽ/, /ʈ/, /ɖ/, final R). For the Norwegians, there were mild varieties both from person to person, but also within the individual’s speaking, like not always uttering /ɽ/ where there was a possibility to use it, which is to be expected. The target

phonemes were then annotated according to how they were actually pronounced, with some extra annotation on the side if there was uncertainty or extra notes to do.

A note to be made on the annotation of “correct” and “not correct” should be made in that there is not always one clear answer to a pronunciation being “correct” or “not correct”, because of the naturally occurring variations within speakers and among speakers. In this case, for all the recorded participants (native and non-native), “not correct” was annotated places where there was a definite deviation from native speech, like not pronouncing a final R at the end of a word appearing sentence final, or pronouncing /ʉ/ as /u/. “Correct” could still be assigned to the deletion of a final R, if the non-native speech matched native speech in places where it would be natural to not pronounce the final R, like in front of a word starting with a consonant. In some cases like these, “not correct” was still annotated, whenever there was a considerable pause between two words. Since there could not be an exact native equivalent with the exact same amount of pause, this had to be judged by instinct as a native speaker.

3.3.5 Perceptual rating on accentedness and comprehensibility by native raters

Three native Norwegians were selected to be a part of the second stage of data analysis. To make sure there were no biases, the raters were not familiar with the recorded native Norwegians.

They were all gathered in the same room on the same day to do the rating session. Each brought their own laptop PC, which all included the possibility to use Excel sheets.
First, they were introduced to the terms “accentedness” and “comprehensibility”, with Norwegian examples played for them, presented on a likert scale from 1 to 9. Then there was a training session, which lasted about 20 minutes. The participants got to hear some examples, which were randomly cut sentences from the participants’ recordings, which again were excluded from the proper rating segment of the session. I asked what they would rate the examples as in terms of accentedness and comprehensibility. The ratings were from 1-9, which I had explained in advance how they should utilise the numbers. The criteria for the rating scales were as follows:

- For comprehensibility: 1 = not understandable at all; 9 = completely understandable
    - 1-3: hard, barely understandable speech; 1 would mean unable to catch any word from the speech.

    - 4-6: middle, more or less understand
    - 7-9: easy, very understandable speech; 9 would mean understanding each word clearly

- For accentedness: 1 = no accent at all; 9 = totally foreign accented
    - 1-3: native to near native; native speaker or near native speaker; 1 would mean this speech is definitely spoken by a native speaker.
    - 4-6: middle: Non-native, accent is acceptable (origin of speaker not obvious)
    - 7-9: foreign accent: accent is heavy and obvious (origin of speaker might be possible to hear). 9 would mean this speech does not sound like Norwegian at all, regardless how much I understood.

We discussed if there were any differences in personal thoughts or definitions of "accent" and what made some think different than the other, until we found a common ground, ensuring that they would judge the accentedness and comprehensibility similarly.

Then the formal rating session began. This included no discussion or communication with one another, and all were to rate on their own. The raters had no access to the recordings other than listening to it from my computer, and the raters did not know about any of the speakers' identities. I had sent the premade rating sheet to each participant, and made sure along the way that everyone was following the correct order. I played the recordings in randomised order, mixing between Danish and Norwegian natives, playing mostly once, sometimes twice if they needed a repetition. After the session was done, the rating sheets were sent to me and gathered into one excel sheet for comparison, which was available to me only.

4. Results and analysis

In this section, I will go over the results of the data collected. First, in 4.1 I will present the accuracy of the acoustic analysis, shown in the tables that will ensue. As mentioned, the results are from annotations done in Praat, marked either as a "present/correct" or "not present/correct". One thing to note is that "present" or "correct" doesn't necessarily always overlap. An example, which frequently happened in this study, is the deletion of final R if the next word starts with a consonant. Naturally, this may also depend on how rapid the speech is, so it would be more

natural to hear the final R if there is a noticeable pause, or the speech is slower. In a similar vein, the retroflex tap/flap, which is a phoneme that does indeed exist in the Urban East Norwegian (and some other dialects), is not a necessity to pronounce and the use or lack thereof does not make you sound more or less "proper Norwegian" (although you would maybe have some speakers of the eastern Oslo area on your neck if you pronounced "Akerselva" in any different way than /ˌɑːkəʂˈʔæɽʊɑ/). The accuracy, or "correctness", was therefore marked based partially on the context, not only if it was present or not. In 4.2 I will go into accentedness and comprehensibility ratings done by native Norwegian speakers, where I will look into any noticeable features, and how this might fit together with the accuracy rate. 4.3 will include the students' answers to the questionnaire done at the end of each session, and in 4.4 I will include a summary of the longer interviews.

4.1 Accuracy rate of target sounds

Here I present the results of the accuracy rate of all phonemes, divided into the three participants (shown on the left side), the two conditions (top row), and sessions (S1, S2 and S3). On the right hand side are the native Norwegians' accuracy rate presented, divided into action and non-action, but not sessions, since they only had one session each.

Table 8

Accuracy rate results

		L2						L1	
		Action			Non-action			Action	Non-Action
		S1	**S2**	**S3**	**S1**	**S2**	**S3**	N1, N2, N3	
Thyra	/y/	100%	62,5%	100%	100%	87,5%	100%	100%	100%
	/ʉ/	70%	90%	85,7%	90%	85,7%	70%	100%	100%
	/ʈ/	75%	75%	100%	100%	100%	100%	100%	100%
	/ɖ/	100%	100%	100%	100%	100%	80%	100%	100%

	/ɽ/	0%	0%	0%	0%	12,5%	0%	0%	0%
	/r/	38,8%	66,6%	88,8%	33,3%	72,2%	72,2%	94%	100%
	/y/	100%	100%	88,8%	100%	100%	100%	100%	100%
Margret	/ʉ/	80%	100%	100%	80%	100%	90%	100%	100%
	/ʈ/	100%	75%	100%	100%	75%	100%	100%	100%
	/ɖ/	60%	100%	80%	100%	80%	80%	100%	100%
	/ɽ/	0%	0%	0%	0%	0%	0%	62,5%	87,5%
	/r/	55,5%	83,3%	66,6%	44,4%	83,3%	83,3%	100%	100%
Alexandrine	/y/	87,5%	75%	100%	100%	87,5%	100%	100%	100%
	/ʉ/	90%	80%	100%	80%	80%	100%	100%	100%
	/ʈ/	100%	75%	75%	100%	100%	100%	100%	100%
	/ɖ/	80%	100%	60%	100%	100%	80%	100%	100%
	/ɽ/	0%	0%	0%	0%	0%	0%	50%	25%
	/r/	55,5%	66,6%	88,8%	66,6%	77,7%	88,8%	100%	100%

In Table 8 we can see accuracy ratings of each phoneme, for each participant and for each session over time. This is, in other words, the most detailed overview of the results, and other overviews will follow below, to see if there are any interesting or substantial differences. The results show an almost striking randomness, and a wide variety of results depending on person and phoneme, person and progression, and phoneme and progression. We may here look at the different participants and comment on each participant's progression.

4.1.1 Thyra

Thyra has a wide variety of results, with improvements, deterioration and steady accuracy, in both conditions (with action and without action). We can even see some dissonance between the

two conditions over time in one and the same phoneme, like with /ʉ/ *decreasing* in accuracy over time when done without action, but first improves and then deteriorates with action. /y/ has an odd "dip" in accuracy on both the conditions, but to various degrees. This particular case could in itself be explained by such natural happenings as her being less tentative to her pronunciation that day, but this pattern does not show with the other phonemes. The phoneme /ɖ/ has a remarkable high accuracy all the way, except for one unexplainable instance. She is also the only of all the participants that did actually pronounce the retroflex tap at one instance in one word during one session, and then to never pronounce it again. If this was purposefully pronounced or not is unknown. The only instance where we see a more apparent improvement is with final R, although the improvement does not show any correlation between the conditions, in that action does not give more improvement than non-action. The only conclusion from this particular example is that the accuracy has improved.

Although conclusive evidence is hard in this case, the data suggests that many of these phonemes are already well incorporated into her language learning, albeit not perfect (this can definitely be said about native speakers as well, but possibly in different ways than being unable to pronounce a certain sound). It seems to require some amount of effort, but considering the language learning has been going on for some time, the effort could possibly be less "forced" over time. I will go further into this aspect later on, when getting into the interviews from the students. The phoneme she struggles the most with is obviously the retroflex flap, and there is no sign of improvement, but as mentioned, this is expected because of several factors, like it not existing in Danish, it not being a mandatory phoneme, and there is no active effort put on her to learn this phoneme. The only place there seems to be some improvement is with final R, which was often observed to be of a difficult task, which is also expected, considering Danish being infamous for "deleting" their final phonemes, or rather that final R in Danish is not a stranger to reducing the final R along a vowel reduction (Basbøll, 2005, 53).

4.1.2 Margret

Margret, shows similar results in that improvement, deterioration and steady accuracy seem random. Both vowels have a high percentage of accuracy, but with some unexplainable "dips", before rising again, while sometimes starting out well, then coming up to 100%, then down again in one condition, but not in the other. The retroflex /ʈ/ has the exact same progression in both conditions, while /ɖ/ has a progression that's hard to understand, with very different results

in the first session, then improvement in the action condition and deterioration in the non-action condition, before deteriorating again in the action condition, ending up at 80% in both conditions. Our dreaded retroflex flap is yet again non-existent. Final R is also again showing some clearer improvement, minus some unknown reason for a less precise result in the action condition. The only clear improvement is from S1 to S2, and afterwards, progression either stagnates, or does down.

Conclusions are again hard to find, but improvement is minimal, if there at all. However, in most cases, in both conditions, the final session shows a higher accuracy than the first or equally good. The deterioration we can see scattered about are hard to explain, considering they do not match the other condition from the same session. The results seem random enough that with Margret, like Thyra, the skills in pronunciation are already high, and explanations for the varied results can be various amounts of effort and energy put into the individual sessions and sentences. There is also the possibility, as the action condition has no apparent positive effect whatsoever, that the action condition in this experiment is inhibiting rather than aiding in easing the cognitive load. I will also get back to this later on.

4.1.3 Alexandrine

Finally we'll look into Alexandrine, which has no radically different results than the previous two participants. In her case, there is more suggestion of improvement on the vowels, generally improving over time. There is an unexplainable difference in /y/, where the action condition starts off weaker than in the non-action condition, and then in both conditions there is a decrease, before both ending up with a perfect 100%. The second vowel also has a random progression independent of which condition, but ends up with better results than the first session. Both the first retroflexes, minus the retroflex flap, which again remains non-present, oddly enough seem to either deteriorate over time in both conditions, or stay the same. For all the students, these have at earlier times been subject to difficulty and sometimes even heavy exaggeration, but considering the two other students have seemingly minor improvements or stay generally in high accuracy percentage, this is somewhat unexpected. This is again paired with the fact that her final R has a very clear, steady improvement over time. However, as with the others, there is no sign of greater improvement in the action condition than with the non-action condition.

We can see that more often than not, the final session shows better accuracy than the first one, but not uniformly. There is some possibility that improvement in one area means less attention and practice given to another phoneme, hence one phoneme improving, and the other deteriorating. There is, yet again, not enough data to give proper conclusions to this alone, but below, I have included different representations of the progression over time, which we will look into now.

4.1.4 Norwegians

From the grid/graph it is clear that the native speech is more accurate than the non-native speech. All Norwegians got 100% in accuracy on all vowels, as well as two the retroflexes /ʈ/ and /ɖ/. Final R also received 100% accuracy for all, except in one instance for N1, which pulled the accuracy down to 94%. This does not mean N1 has less accurate pronunciation overall, it could simply be a single instance of a minor pronunciation error, which is happen for native speakers as well. The results for the retroflex /ɽ/ is to be expected, by this phoneme being a variation of pronouncing L (and sometimes R, not included in this experiment), and so not all native Norwegians (nor speakers of Urban East Norwegian) would pronounce the /ɽ/ at all times. Variations interindividually is also to be expected. In this sense, the Norwegians display an expected high accuracy of all phonemes, with the natural variations.

Further, in Table 9 we can see all the mean accuracy rate of the phonemes for each participant at each session, divided into each session and the two conditions. It also includes the accuracy rate for all participants and all phonemes, giving an overall accuracy rate. Here, the Norwegians' accuracy rate is not included.

Table 9

Accuracy rate of all target phonemes, and accuracy rate of all phonemes for all participants

	Action			Non-Action		
	S1	S2	S3	S1	S2	S3
Thyra	56,6%	64,1%	79,2%	60,3%	71,6%	67,9%
Margret	62,2%	77,3%	69,8%	62,2%	75,4%	75,4%
Alexandrine	64,1%	64,1%	75,4%	69,8%	71,6%	81,1%

Action			Non-Action		
S1	S2	S3	S1	S2	S3
61%	65,5%	74,8%	64,1%	72,9%	74,8%

The results of the upper part of Table 9 indicate a somewhat steady improvement over time, but only in Thyra's case there is a minor indication of favourable improvement in the action condition than the non-action condition.

In addition, I took a look at some individual words and their statistics. Initially, I had an idea that stronger emotional words like "hate" could spark some embodied enthusiasm and therefore trigger the final R to be more heavily or easily pronounced, especially with action, and there seems to be some truth to it, since the final R was more often pronounced than not (and in two cases, action seemed to trigger pronunciation of this final R with Thyra in S1, Alexandrine in S2). Considering that this is a disyllabic word appearing before a word beginning with a consonant (the type of context in which the final R seemed to be the hardest for the participants to pronounce), this could potentially be interesting to look further into. However, no conclusion in itself can be taken from this one particular word.

There is also a strong indication that pausing at the end of a disyllabic word with a final R gives room for better pronunciation, which is not surprising, given that they can put more effort into the word rather than "rush through" the word when it appears mid-sentence. There are also signs of minor improvement specifically on the disyllabic word "eier", which always came before a monosyllabic word starting with a *vowel*, which should give room for the pronunciation of the R. First session showed that none pronounced the final R in this word, second session showed one instance of pronunciation, but without action, and the third session had the most instances, with one participant pronouncing it in both conditions, one participant not pronouncing it in either, and one participant pronouncing it with action. Similarly, this in itself is not enough to give conclusive evidence of improvement, but there are signs that suggest minor improvements, individually, and somewhat sporadically, over time.

As for the overall accuracy rate for all participants, in this particular case, they seem to end up doing equally well in both conditions. The non-action condition seems to do better in both the first two sessions, and is finally caught up by the action condition in S3. This could indicate that the action condition is not entirely ineffective, but, as I will explore in the discussion, the action condition does not seem to have any immediate effect momentarily, or spontaneously.

4.2 Accentedness and comprehensibility rating

In Table 10 are the presented results from the rating session. The table is divided into Accentedness (ACC) and Comprehensibility (COMP) on the very left, with the respective ratings belonging to it, and into the two conditions of Action and Non-Action at the top. Participants' individual results are shown horizontally next to the respected name. Each number represents the mean rating for each recording session (S1, S2, S3). To the right is the L1 speaker rating, also divided into Action and Non-Action, and into individual Norwegian participants (N1, N2, N3), but not session, since all Norwegians had only one session each.

Table 10

Accentedness and comprehensibility rating results

		L2						L1		
		Action			Non-action			Action		Non-action
		S1	S2	S3	S1	S2	S3			
ACC	Thyra	4,9	4,9	4,5	4,2	4,04	4,1	N1	1,01	1,04
	Margret	4,1	3,75	4,6	4,6	4,4	4,9	N2	1,01	1
	Alexandrine	4,9	5,1	4,3	4,8	4,6	4,1	N3	1,2	-
COMP	Thyra	7,1	7,3	7,6	8,2	8	7,67	N1	8,75	8,9
	Margret	7,1	8	7,45	7,6	7,9	7,9	N2	8,58	8,85
	Alexandrine	7,4	6,7	7,6	7,15	7,19	7,9	N3	8,75	-

Here is presented the results from the final data collection session, which included three native Norwegians (who were unfamiliar to all the voices recorded) listening to the randomized recordings and rating them accordingly to the agreed rating measurements. One of the Norwegian recordings was taken out of the rating session, because of time issues, and is left blank. The left hand side shows the ratings divided into two categories, accentedness (ACC), and comprehensibility (COMP). S1, S2 and S3 stands for Session 1, Session 2, and Session 3. On the right hand side the ratings for the native Norwegians are shown, marked vertically by the coded names N1, N2 and N3 (N standing for Norwegian).

First we can take accentedness into account. All of the L2 learners' accents were rated generally as mid-accent, that is, the native Norwegians would judge that the speech was non-native, but it was not heavy. Surprisingly, there is no clear pattern of improvement of accent (reminding that: lower number = less accent). In some cases, the accent even shows signs of being *heavier* than initially. Why this could happen is uncertain. The native Norwegians were never instructed to

look for certain sounds or “mistakes”, nor certain things that made the accent good or “passable” in any way. One apparent mistake many learners of Norwegian make, including these Danish learners, is the mixing of tonemes. To learn the tonemes can be very hard, but even after learning to distinguish them, knowing when and where to use the right one is a matter of almost complete guessing to the learners. This in turn, being a very pronounced (no pun intended) part of the Norwegian language, could likely make a native Norwegian rate accent as heavier if the toneme is wrong. The native Norwegians would no doubt notice these kinds of errors, and could potentially make the rating of a whole phrase/sentence somewhat skewed towards more accented, even if the rest of the pronunciation could be considered close to perfect, or very good. The native Norwegians were not asked post-rating session of specific ratings or thoughts on specific errors or pronunciations, so the exact reason for their individual ratings will remain unknown.

The dissonance between the action condition and the non-action condition is sometimes striking, with progression over time in individual participants sometimes going in opposite directions, like with Alexandrine, or that the progression shows no great sign of improvement other than minor decimals, like Margret. There is also no sign of the action condition giving less accent impressions than the non-action condition. The occasional increase in accent rating also does not seem to correlate with the accuracy percentage, where the students generally, but only slightly, improve over time. Comprehensibility seems to generally improve *somewhat* over time, except for Thyra in the non-action condition, where it only decreases. The improvement is nevertheless very minor. As predicted, the native speech is considered very low in accent, and high in comprehensibility. Perfect comprehensibility is also not expected, since to be comprehensible doesn’t necessarily entail accent (Galante & Thomson, 2017).

The raters were rating blindly, because the recordings were put in randomized order, this to validate the rating. By this, the raters would judge the L1 speech based solely on their own native judgment. They could clearly hear which of the recordings were spoken by a native Norwegian and which were not, giving L1 speakers high comprehensibility scores and low accentedness scores. This tells us that the native raters were sensitive enough in their listening, and clearly understood the assignment. This identifies the raters as equally able to rate the L2 speakers as non-native and we can confirm their ratings as legitimate.

Table 11 presents the mean value of ratings for all participants.

Table 11

Accuracy and Comprehensibility Rate Results For All Participants (Including L1 Speakers)

	L2						L1	
	Action			Non-action			Action	Non-action
	S1	S2	S3	S1	S2	S3		
ACC	4,6	4,59	4,5	4,59	4,7	4,4	1,1	1,02
COMP	7,2	7,39	7,5	7,66	7,7	7,8	8,7	8,9

We see that there is an improvement of accent over time, but the changes are very minor, especially in the case of the action condition. This tells us that if we look at the results for each individual, the accent rating goes up and down, with a certain amount of improvement. However, we can see, despite some individual instances pulling the accent score up, the numbers point in a favourable direction. That said, the decrease of accent over time as an average between three individuals, are very minor. The action condition and the non-action condition ends up very similar in rating at the final session, proving there are minor differences.

In terms of comprehensibility, there is some improvement over time. In this case as well, the non-action condition ends up with better results than the action condition. From both of these results we can deduce that accent and comprehensibility has a certain improvement over time, even if very minor, but adding action either has no effect, or even prohibits some momentary improvement, whether in this case if it is acoustic improvement or perceived improvement. Comprehensibility scores are quite high for both conditions, showing that comprehensibility and accentedness, although related, can be separated as not dependent on each other. Accentedness

scores tells that the raters found the accents to be “definitely present”, but comprehensibility scores are closer to the native scores in comprehensibility than the accentedness scores are to the native scores in accentedness.

4.3 Questionnaire

Here, we can see the answers the students have to the questionnaire that they were given at the end of each recording. The questions were the same each time. This doesn’t necessarily give room for a wide variety of answers, but we can see if there is some consistency. The questions in their entirety can be seen in the appendices. The answers the students gave are here made into a scale from 1 to 5, 1 symbolizing the weakest grade answer in relation to the question, 5 being the strongest.

Table 12

Questionnaire answers

	S1			S2			S3		
	D1	D2	D3	D1	D2	D3	D1	D2	D3
How often doing action when practicing?	4	2	5	4	5	5	5	5	4
Are you comfortable with actions?	5	5	5	5	5	5	5	5	5
What degree does interacting help?	5	4	5	5	4	5	5	5	5
How often using Norwegian outside of school?	5	4	5	5	5	5	5	5	5
What degree does actions help?	5	4	4	5	4	4	5	4	4

The answers show that the students enjoy these kinds of exercises in their language class. There is one curious instance in the first session where Margret, rates the frequency of action during practice as very low, while later on, it is rated as very high. This could imply that the definition of “action” has not been communicated well enough, since the frequency of the use of action in classes neither has gone down nor up before or after the first session. What seems to have happened is that the student understood the question differently in session 2 and 3, while the two other students had possibly a similar idea of what “action” meant from the start. But the question if they did actually understand what the relevant definition of “action” was might be asked here. I will address this in the discussion chapter.

Finally, we can deduce that there is no correlation between the answers to the questionnaire and the results from the accuracy rate and the accent and comprehensibility rating. The students

marks “action” as helpful, enjoyable, and of frequent use. In addition, they report interacting with each other, the teacher and other students, and that they very frequently use Norwegian outside of school. This last factor in itself may very well contribute greatly to their Norwegian skills. We will also get back to this in the discussion.

4.4 Interview

This section will talk about the interview done post recording of the students action and non-action speech. The interviews were recorded and then roughly transcribed into the most important and interesting remarks done by the participants. These remarks and thoughts will be compared to the results of accuracy and ratings. The interviews will be written in sections with focus on one participant at a time and their own respective progression through the questions. I will also present the questions from the interview, which will also be included in Appendix C.

Table 13

Interview questions

Norwegian	English
Hva synes du er vanskeligst med å lære norsk?	What do you think is the hardest part of learning Norwegian?
Hva synes du hjelper mest for å lære norsk generelt?	What do you feel, generally, helps you the most in learning Norwegian?
Hva slags aktiviteter bruker dere i språktimene? Beskriv dem	What kind of activities do you do during language class? Describe them
Hva synes du hjelper mest for å lære norsk i språktimene?	What do you feel helps you the most in learning Norwegian in language class?

For Thyra, the answer to what was the hardest part about learning Norwegian, the recurring answer was tonemes, often mentioning the reason being Danish does not have this feature (Basbøll, 2005; Kristoffersen, 2000). Thyra was the only one of the three students who reported she had been living in Sweden for some time prior to moving to Norway, and understood and spoke Swedish (not natively). Swedish, being another European and Scandinavian language that does have tonemes (albeit different than Norwegian) (Andersson, 1994, 274-275), as well as

sharing some phonemes with Norwegian that Danish does not have, might have given Thyra some advantage when learning Norwegian. Nevertheless, she found tonemes to be difficult overall, telling that, despite her previous knowledge of tonemes, Norwegian and Swedish had different patterns in tonemes. In comparison to Norwegian, Danish has *stød,* suprasegmental phenomenon related to phonation and accent (Basøll, 2005, 83). The issues of tonemes also felt present to her during performance on stage. At the final interview she expressed the difficulty of not "getting lazy" with her language learning, meaning that at that point, she was being perfectly understood and comprehensible. At this point, the language classes had ended, meaning she would need to keep up her own language learning by her own work force. This does somewhat correspond to a remark she gave, where she was aware of the very steep learning curve in the beginning when learning Norwegian, which then quickly started to stagnate, with less progression over time when all the students started getting experienced and familiar with the language, and they were considered "fluent enough" by themselves and others.

The most useful techniques in everyday life to learning Norwegian to her was mere "usage" of the language, especially to other Norwegians, as well as repeating single words and sentences over and over. This also included reading and listening to Norwegian. The only people she spoke Danish to were the other Danish students. Within language classes, she felt the mimicking practices helped a lot, which included listening to a native Norwegian having recorded speech, which the students then either mimicked by miming, humming the melody, or directly copying the speech with all its traits after listening. These classes involved little acting technique, in the sense that they put meaning and emotion into the sentences they were practicing, as opposed to the classes I observed. The repetition during class, she said in the final interview, also helped "training the tongue" and seemingly helped the sounds that felt "unnatural" feel "natural" after a while. She had also found her own technique during one language class, which included writing and drawing, one including writing the words as she heard them herself, and one where she drew lines following the melody of the speech.

One of the questions asked the participants to describe the activities in class, to get an idea of their most memorable exercises, and what they thought about them. One remark she had on the language classes I had been observing was the almost "childlike" way of learning to speak, mentioning the minimal focus on grammar, and rather on learning to speak again. She favoured this way of learning, instead of focusing on "the theoretical" as she called it, referring to a more typical grammar class. Another exercise she enjoyed was during the other language class, an

exercise she called “pendling”, which included exchanging between pronouncing a sentence in their native language and then in Norwegian, to feel the difference. This, she said, made her more aware of the differences between the languages.

For Margret, the similarities between the two languages proved to be a challenge, especially considering the often identical vocabulary, but that she had to pronounce it in “bad Danish”, and relearn a new way to pronounce familiar words. She also expressed surprise over the difference in how the two languages pronounced the alphabet, meaning the students had to learn the alphabet all over again, echoing what Thyra said about the almost “childlike” feeling of learning language all over again. One of her greatest difficulties were the rolling R’s (NOR: *Rulle-R*). Later on, she said she was doing the rolling R’s, but were “overcompensating”, meaning she pronounced them too harshly, presumably more like a trill /r/ than a flap/tap /ɾ/. This shows that she worked hard on this phoneme, ending up over-emphasizing the phoneme, similar to what happened in Estrada’s (2004) study, but later on trying to tone it down again. She also expressed difficulties with tonemes, and not being able to hear the difference in the more “subtle examples” of words. Interestingly, Margret was aware of the de-voicing phenomenon of Danish (Basbøll, 2005), and the use of voiced consonants in speech proved to be difficult, by her not being used to it.

Margret did express “the integration of stage performing” (NOR: *scenisk integrering*) as somewhat difficult, though not as a hindrance to her flow in speech. She talked about the apparent difference between speaking Norwegian privately with friends, and speaking from a script on stage while considering other factors like body language, where to stand and walk. In the latter cases, she felt she no longer spoke in her “natural Norwegian voice”. At the same time, not being a direct hindrance in her speech, she said acting on stage also helped in “playing more with melodies” in the language, to “go to the edges of speech sounds”, and that she dared and allowed herself a larger variety of speech. Similarly, during language class, putting meaning into the text gave her the courage to “throw herself out into the text” and not be too hung up on how to pronounce the words in exact detail, which she felt made her speech more clear.

«[The acting on stage] doesn’t impede the language itself, but it makes it easier in many ways, because one allows oneself a bit more to play with the melody, and go into the edges of the language sounds. But I also really rise in pitch. My voice gets much lighter on stage. […] I don’t

do that to the same degree [in Danish]. […] And I have a somewhat powerful voice, and that disappears quite often in Norwegian, unless I'm very aware of it»

- Margret, 15.12.2022

Finally, she expressed the feeling that she somewhat "changed her personality a bit" when speaking a foreign language (this also included English). This shows that acting includes aiding elements, like the flow of speech, but also impedes some of the language learning, like less opportunities to concentrate on pronunciation, and maybe an "othering" feeling when speaking your non-native language. I will get back to these issues in the summary and the discussion in Chapter 5.

Similarly to Thyra, Margret expressed the issue of Norwegians understanding her very well, omitting the immediate need to learn the language and "perfect it". This would then lead to less corrections and less self-awareness of her pronunciation. Another issue she had from the beginning of the interviews was putting the word stress on the wrong word when speaking (NOR: *feil vekt på meningsbærende ord*), especially on stage. She later said this could be because of Danish *stød*, being used to putting *stød* on certain words, which then possibly could be erroneously "translated" into Norwegian as word stress.

Margret reported at the second interview that she had started speaking Norwegian with her partner, which she felt helped a lot on her speech, vocabulary and pronunciation. At that point, the only time she spoke Danish was with her family, and the other students. This could be a contribution to making her L2 a more "personal" part of her life, contributing to positive feelings. If this in turn would have a positive impact on accentedness is not clear, though not impossible (Galante & Thomson, 2017).

Finally, as with the other two participants, Alexandrine found tonemes to be the hardest part about Norwegian, noting the lack of tonemes in Danish. It was also a part of Norwegian pronunciation one "just had to know about". She also found the rolling R's difficult, echoing Margret's difficulties. Alexandrine also expressed difficulties hearing the difference between /ʉ/ and /y/, feeling they were more similar to each other than the Danish equivalents. She could here be referring to the Danish /ʊ/ and /y/, which then would be in agreement with her statement of the Norwegian vowels feeling "too similar" in proximity to each other.

She reported that the greatest aid in learning Norwegian outside of class was continuously speaking it, trying to create "muscle memory", and often mentioning music and singing as a great aid, reporting she had been told she sounded "more Norwegian" when singing. In one instance she mentioned a day at school which involved singing throughout the whole day, reporting that at the end of the day, her voice felt "warm" and her "speech flow" felt good. Although music and musicality is not part of this thesis, this is an interesting point of study, which several researchers has tied to pronunciation studies (Delogu & Zheng, 2020; Nardo & Reiterer, 2009; Zhang et al., 2018; Li et al. 2020; Chobert & Besson, 2013). During class, she felt the balance between focusing on segmental phonemes, putting them in difficult tongue twister-like sentences, and putting meaning into them all contributed well to her development. A technique she referred to as "blocking", involving obstructing consonants by not opening the lips while pronouncing Norwegian obstruents like /b, d, g/, which presumably helped the students practice voicing of consonants. Other exercises she found helpful in her speech were any that helped her feel relaxed and then gave room to "give meaning to the words", use background music to help put emotion and "extremities" into her speech, and spend more time learning and "feeling" the script and text, and exercises involving "forankring", roughly meaning "to anchor" the words, which often involved tying different kind of physical movements and involving action with the speech. This is interesting in the sense that this describes very well one of many ways EC can be incorporated into language learning and pronunciation practice.

In sum,the students found tonemes, rolling R's, and the lack of sense of "emergency to learn the language" and the stagnation of improvement. To be aware of the differences between two very similar languages proved to be somewhat of a challenge, and the less rule-driven differences, like tonemes vs. *stød*, were of the greatest difficulties. They found many different kinds of action to be of aid, including "adding meaning", "blocking", mimicking native speech, and writing the melody. In everyday life, to continue speaking, getting corrected or asking for correction, as well as speaking it in private and personal settings were reported to help a lot.

4.5 Summary

The analysis of development in the students' individual accent in their L2 indicate individual patterns and variable benefits from using action/acting as a condition on a sentence involving their usual practice phonemes. Overall the students all together were somewhat improving in their accent precision over time, but not in any radical way. Individually, there are signs of both

improvement, and some setbacks, sometimes with a surprising difference in development in the two different conditions over time, where one condition might improve more than the other from one session to the other, or even one condition improving while the other deteriorates. Nevertheless, on individual levels but globally for all target phonemes gathered, they end up better than they started.

This does indicate that the learning curve with learners of high skill and experience in the target language is somewhat bumpy. There is improvement, but it is slow, and not without some decrease in precision, followed by improved precision, and vice versa. Also for all participants, the precision of individual phonemes has less of a steady improvement, but rather a progression where improvement and decreasing precision, as well as a static progression, all happen interchangeably, sometimes even decreasing in one condition while improving in the other. The only phoneme that showed a somewhat steady improvement over time in all participants and in both conditions was the final R.

There is little sign of action having a direct positive or improving effect on the pronunciation. Action in this study was used momentarily and as a tool to possibly lure out some precision or "naturalness" that could trigger better pronunciation, which did not happen. There is also no indication that using action prohibits improvement over time. However, the use of action in this case had no obvious beneficiary effect over non-action; possible explanations for this will be taken up in the discussion.

The ratings for the learners were generally favourable, they had very high comprehensibility scores, and accentedness was mostly in the area of "medium accent". Overall, accentedness did decrease only somewhat over time, and comprehensibility increased with a few decimals more. The accent with action went down 0,1 points, accent without action went down 0,19 points, comprehensibility with action 0,3 points and comprehensibility without action went up 0,14. Interestingly, comprehensibility with action had the highest improvement in terms of numbers. There is no obvious reason why this might be. The change is quite minor so it is hard to say if the raters heard better individual words, or if the overall message seemed clearer. The latter is also hard to know since the raters heard the same sentences repeatedly during the session. If we focus on accent, the ratings are random, with no indication of an obvious improvement over time, even though in some cases end up better than they started, but the other condition on the

same person might end up worse than initially. There is also no indication here that the action condition gives the impression of better accent, not in the moment, nor over time. At the same time, there's no indication that the action condition makes them deteriorate over time.

Although the results show that action condition does not directly influence accentedness nor comprehensibility with a stronger improvement than non-action, we cannot completely rule out that action contributes to the improvement, based on the improvement over the three sessions, and the positive attitudes of the students. They favour the use of action and utilise it often in class. In the interviews, the students express difficulties with tonemes and the rolling of R's, and often the issue of keeping the practice up, and to get motivated enough to continue to improve. The process is rapid in the beginning, but after a while, it stagnates, and audible improvement is harder to be aware of. To use the L2 in private, everyday life is reported to be of help, along with assisted learning involving detailed work on sounds, phonemes, stress, melody, and so on, and putting meaning into the words and phrases they use during class.

The final data indicates that performing actions simultaneously with speech may hinder some pronunciation accuracy. It might be that the cognitive load becomes too heavy when they are being recorded while adding emotion and including minor bodily movements. This, however, does not falsify the role of action training in L2 pronunciation classrooms. I will discuss this further in the discussion.

5. Discussion

In the discussion I bring together the findings from the project and discuss what the implications of the results are, how this relates to the research questions, and what the possible faults and improvements of the project could be. **5.1** will give a brief summary of the study, why and how it was done, and what was found. Section **5.2** will focus on answering the research questions based on the results. Section **5.3** will be dedicated to discussing the importance of action on the accuracy of sounds, accentedness and comprehensibility, related to the effects found in this study. In section **5.4** I will make some remarks on EC, and tie in the relation to teaching Norwegian, and suggest some future research in **5.5**. In **5.6** I will discuss some limitations of the design of the study, and possible improvements.

5.1 A brief summary

This project consisted of a short quasi-experiment involving three individuals learning a language that is considered to be very close to their native language. Norwegian and Danish are two mutually intelligible languages due to historical and linguistic reasons, making these two languages an excellent point of interest. Regardless of the definition, learning the other requires still quite a bit of work, with some new grammar involved, and learning a new phonetic system, which includes both "throwing away" certain familiar sounds, and including some novel ones. To learn a new way of speaking, and then to attempt to "perfect" the new way of speaking seems to be an ever interesting topic, not only to linguists, but also to those who actively work with speech, pronunciation, meaning and words in their craft, namely actors. I found there to be an opportunity to study how one artistic profession could fit into the world of linguistics, and the study of phonetics, language learning, pronunciation, and finally the scientific field of EC.

I went into the collaboration with the acting institution these students went to, and I got a direct view into how they work with teaching non-native Norwegians to get familiar with the sounds; the vowels and consonants, the prosody and the tones. This way, I was able to design a research project, attempting to document some possible changes and improvements in their pronunciation over time, after they had already learned quite a bit of Norwegian, and even started speaking it in everyday life. There had been many studies about gesture and similar tactile movements in relation to improved pronunciation, and studies involving the use of drama in the classroom. But none had looked at acting students, and the unique situation of a language relationship like Danish and Norwegian, which gives room for rapid learning, not needing to focus on grammar as much as learning to pronounce the L2, gave the opportunity to see if the use of acting, done by L2 learners who also were comfortable with acting, in the moment could have an effect on pronunciation and accent. This was redefined as "action" in the project, with "non-action" being its contrast, working similarly to a "neutral" stance, or a "less engaged" way of pronouncing. The students were given sentences to read out loud by me, designed to make them pronounce the phonemes they had already been working on, without attempting to make it too clear exactly what was being looked into. These sessions were recorded and analyzed, looking for precision and accuracy in the individual target phonemes. Later on, they were judged blindly and somewhat randomly by native speakers.

The findings indicate that action as a tool used simultaneously with speech in this case did not give any favourable results over non-action, other than a general, minor improvement over time. Individual phonemes did not improve smoothly over time, but rather had some ups and downs in their accuracy. The only phoneme that had a relatively steady improvement was the final R, a phoneme in certain positions and done with a certain speed that they had struggled with. They did report to enjoy using "acting" and "action", though whether the actual definition was ever truly disclosed or communicated well enough by me as a researcher is somewhat unclear. During the questionnaire, I introduced the term "action" (NOR: *handling*) as something physical, an actual action, a gesture, and smaller motions, as well as putting in emotions, subtext (NOR: *undertekst*), or explicit meaning in the text. I had been introduced to some terms they knew through their acting classes, which I also utilised when asking questions and handling the questionnaire. Note though that the questionnaire and interview always happened after the pronunciation recording, and there is a chance that they didn't get a clear enough reminder of what "action" could be. However, during recordings, there was a clear distinction between the use of "action" and "non-action", so some understanding was apparent. There is a possibility a different term could have helped further, but it is hard to say. There is also a possibility that using the actual term "acting" instead of "action" could have helped them, or even adding some "character development" to add to the immersion. I will get back to possible changes that could have helped further in this chapter.

5.2 Answering the research questions

In this thesis, I seek to ask and answer two research questions. Research question 1 asks: Is there a substantial or noticeable effect of the use of action vs. non-action on the pronunciation accuracy of nonnative sounds?

The thesis has been looking at previous studies where language learning and drama-based approaches in teaching are combined as a way to improve language learning, and particularly how it can affect accent. The information I was seeking was to see if this kind of embodied action could lighten the cognitive load of pronouncing somewhat novel phonemes, and thereby aiding better pronunciation and lessen the accent. Based on the findings, this seems not to be the case. The action condition did not yield any favourable results compared to the non-action condition, and when it did, it seemed more random than not.

In the classes where the students went to to learn Norwegian, they received mainly explicit instructions, with a lot of focus on the suprasegmental sounds, before putting them in larger segments, both simple sentences, as well as "tongue twisters" involving many instances of a particular phoneme to practice (not necessarily tongue twisters for a native, but for the L2 learners, this would often be a challenge). In the recording sessions, they received no instruction on pronunciation. Their only instruction was to read the sentences "neutrally" and "with emotion" (NOR: *med innlevelse*), so they received no help during the recordings. However, the students all came directly from their private language class, meaning they either had just received input or were likely mentally prepared to start practicing. In Galante & Thomson's study (2017), they found that the mix of both explicit learner-centered instruction as well as drama-based intervention improved oral fluency. However, although all students received lower accent ratings over time, there was no significant difference in accent rating between the control group and drama group. In this study, there was no control group with Danish students that did not receive drama-related instruction, or that only received a more typical classroom instruction. There is clearly little difference between the different conditions within the individual student, which would imply that the instruction very well may be of aid in accentedness and comprehensibility, but different conditions like action and non-action does not give notable difference in improvement. The lack of a control group of other Danes makes it hard to conclusively say if the instructional methods of explicit instruction and teaching, as well as all the language classes involving acting/action and physical exercises is of higher gain than solely explicit instruction. This could be of further interest to study. What is more conclusive is that, paired simultaneously with speech, action seems to have no direct positive effect on the pronunciation in terms of accentedness and comprehensibility, although it has no clear negative effect either.

Within the sentences used containing target phonemes, the focus was on first-person written sentences, in an attempt to hopefully better trigger the emotions in the action condition sentences, referring to Galante & Thomson (2017) discussing whether first-person text/speech could yield better results, and Estrada's (2004, 2007) use of the emotion of surprise for her subjects. Results show that there were overall accuracy improvement of both accentedness and comprehensibility long term, with some individual differences, where some L2 learners had higher improvement in percentage points than others, which is to be expected, when naturally not all learners learn at the exact same speed and in the exact same way. The action condition had no clear improvement compared to the non-action condition, though, as just mentioned,

there were individual differences, with some L2 learners having a greater improvement from Session 1 to Session 3 in the action condition than the non-action condition, and vice versa. In most cases, individual phonemes either showed improvement from S1 to S3, or remained at the same percentage from S1 to S3, which often would be phonemes the participants already got 100% score in accuracy. The instances where some deterioration in accuracy was present would not be considered important. The deterioration instances were either followed by improvement in the next session, or had been preceded by better accuracy rates in both sessions. This implies that although action may not have a substantial positive effect over time, compared to non-action, there is no apparent negative effect either. This is further suggested by the overall improvement for all participants, where both conditions yielded an equal accuracy percentage in the final session.

Finally, this study agrees with several of the reports in Abenoja & DeCoursey (2019) of enjoyment tied to the use of drama or action in class. Like in their study, the Danish students reported the use of action in class as enjoyable. A factor that I did not necessarily take into consideration during the study would be that it is not known which of the classes the students would be thinking of when answering the question, the classes I observed, or the bigger classes involving more physical exercises (the question simply referred to their language classes in general, not the specific ones I observed). The interviews shed some light on what parts of the language classes they enjoyed, which does give some indication of their possible thoughts in the questionnaire, but it is still not certain if the questionnaire was clear enough, and it could have been beneficial to specify in the questionnaire. This is a limitation to the design of this study, and should have been taken into consideration, and a closer study of the latter classes would have contributed positively to the study. Nevertheless, positive feelings towards the classes were very present. Reports from students in Abenoja & DeCoursey mentioned playing drama as "[helping to] speak the words correctly and with emotion" (2019, p. 726), and that it was easier to remember the correct pronunciation of the L2 words (2019, p. 729). A direct comparison between this study and Abenoja & DeCoursey (and possibly a large amount of language learning studies) is somewhat difficult, since in their study, the students were Chinese residing in Hong Kong (and likely speaking either Mandarin or Cantonese, or possibly both), receiving instruction in English and learning French, which have a much greater difference cross-linguistically than Danish and Norwegian, since in the latter case, most of the focus is on pronunciation, while grammar and learning new words (or learning the Norwegian meaning of a word that exists in both languages) takes more of a passenger seat role. There could likely be some effect of better

memory of grammar and words for the Danish students, but this was not specifically reported, and cannot be concluded so. Nevertheless, this suggests that frequency and enjoyment of the use of action as a condition when practicing and using L2 doesn't necessarily entail that it aids pronunciation, at least not momentarily.

The second research question asks if there is a substantial or noticeable improvement with the use of action vs. non-action with overall oral proficiency on the measures of comprehensibility and accentedness?

In studies involving EC, like using drama in the classroom, or using gesture to ease or aid cognitive information when learning pronunciation, results show that to use EC and different types of action is favourable in the long term to learning language, with accentedness and comprehensibility improving over time, as well as improving enjoyment in the classroom, and improving fluency and memory of vocabulary. There are, however, few studies showing that accentedness and comprehensibility, compared to control groups, improves notably long term. Similarly, in this study, results of accentedness and comprehensibility rating show a minor improvement long term, though not substantially better than in the non-action condition. Ratings show that individually, one participant received a *higher* accentedness rating over time in both conditions (Margret), while the two other participants received lower accentedness rating over time. In all individual cases, action had no substantial positive effect on accentedness. Comprehensibility scores developed positively from S1 to S3, except for Thyra in the non-action condition. As predicted, native speech is considered very low in accent, and high in comprehensibility. Perfect comprehensibility is also not expected, since to be comprehensible does not necessarily entail accentedness (Galante & Thomson, 2017). Overall for all participants, accentedness received lower rating over time, though in a minor degree, with non-action received 0,1 lower point than action in the final session. Comprehensibility improved in both conditions, and non-accent improved somewhat more. Based on these results, overall oral proficiency does not substantially improve with the use of action nor non-action, since the differences in end session rating are quite minor.

In comparison to many of the studies mentioned in this thesis, the L2 learners in this study were learning a very similar language to their own, and had already gotten to a very decent proficiency, meaning they would be considered experienced speakers. Possible explanations may lie within the SLM-r theory. As has been mentioned, the same mechanisms and processes are

applied both in learning L1 a later on learned L2, but the results of this application are different between the two languages, in that L2 learning is shaped by perceptual biases from L1, making L1 interfere with L2 sounds (Flege, 2021, 24). The L2 learners in this thesis can be considered "late" learners, of which SLM-r focused on by comparing with "early" learners, rather than the experienced bilinguals-focused SLM. SLM-r also abandoned the SLM notion that there would be an upper limit on accuracy with which L2 sounds could be produced, and rather that segmental production and perception coevolve without precedence.

Still, Flege claims that L2 learners can never perfectly match monolingual native speakers of the target L2 (2021, 64). One other possible factor could be that the listeners would tend to assign harsher scores when rating the accent, compared to comprehensibility (Derwing & Munro, 2020, 17). This further suggests that comprehensibility requires less native-like traits to be understood, and to be acceptable, while accentedness is judged with higher requirements. The results in this study also illustrate that an addition of a particular emotion, and the use of action simultaneously with speech would not necessarily be considered an improvement for accentedness by raters, at least not for experienced L2 learners. When compared to natives, the L2 learners clearly have a higher rated accentedness, but also a high comprehensibility rating. This further adds to the previous research suggesting that a very present accent does not entail lower comprehensibility (Galante & Thomson, 2017, Derwing & Munro, 1995). In sum, comprehensibility rating receives less requirements, accentedness rating receives less tolerance for error, and the two are independent dimensions of L2 pronunciation.

Literature on language learning has discussed that it is wiser to stick to the principle of intelligibility rather than the principle of nativeness, since the former is more likely to lead to improvement in comprehensibility (Derwing & Munro, 2009, in Galante & Thomson 2017), calling the goal of nativeness "unrealistic" and "unnecessary". This could be the case in many aspects, since accent, generally speaking, seems to be an eternal "presence" for many, if not most, learners of a language. Flege also talks about mastering L2 sounds, in SLM-r theory, as something to be abandoned, claiming that L2 learners can never perfectly match the native speakers of the L2 target (2021, 64). Within language learning, I agree with this sentiment, since accent is no handicap, when comprehensibility (and intelligibility) is far more important for communication. In terms of acting and the theatre, this is likely to be up to debate, and goes beyond linguistics in terms of questions of inclusion, tradition, and how accent, dialect, and even speech impediments can fit into a narrative (either as an obvious feature of a character, or to be

ignored). The study of "mastering" L2 sounds, either by monolinguals or bi/multilinguals, however, should not be put completely dead.

Finally, there is a possibility that the "othering" feeling, and the "changing of personality" when speaking Norwegian (on and off stage) might have contributed to making the speech feeling less personal, being aware of the "non-nativeness of the language". This possibly contributes somewhat to the pronunciation and flow of speech in L2, by loading the cognitive flow. Although the students' progression in Norwegian was rapid, and their skill level was high, they had spoken the language, at the time of the interviews, for less than two years. Though by making the language more "private", like speaking to a partner, was reported to be of aid. Flege noted that everyday speech involving meaningful and engaging content would be of high quality (and quantity) L2 input, giving room for natural speech flow and social engagement (Flege & Bohn, 2021). High quality L2 input would then, naturally, lead to better learning.

In the next section, I will go into how action can impact accuracy of sounds, accentedness and comprehensibility, and discuss whether the impact might be important or not, as well as discussing if mastering L2 sounds is possible.

5.3 Importance of action on accuracy of sounds, accentedness and comprehensibility

The results in this study stand with some agreement on previous literature on the use of action to aid pronunciation training. The focus of the experiment has been on the role of embodiment in lightening the cognitive load. In this section I will discuss the importance and impact of action. I will discuss *how important* action is on accuracy of sounds, on accent and comprehensibility, and on the *improvement* of sounds, accent and comprehensibility. I will also discuss the relationships between the measurements of accuracy, comprehensibility and sound accuracy.

The studies mentioned in this thesis have often shown that different forms of action have positive effects on sounds, accentedness, comprehensibility, as well as fluency, enjoyment and the realization of vowel length. There is some uncertainty how certain forms of action can have a positive effect on accentedness and comprehensibility, like drama in the classroom. Further, the use of stronger emotions like surprise can aid the pronunciation difficulties of L2 speakers of French, but without clear evidence how this proves effective over time. Most studies also involve

learners of an L2 notably different to their L1, in that grammar, pronunciation and overall structure might be quite different. In some cases the language of instruction is not even in their own native language. Considering this, studies on the role of action in language and pronunciation learning has had a wide variety of results.

In this study, the lightening of the cognitive load did not seem to occur for the students, taken from that the improvement over time was not notably greater for the action condition than the non-action condition. It is possible that the way action as a method was presented in the recording sessions could end up taxing the cognitive load rather than aid it, which is not ideal when speaking. This was an observational study, and the students were not trained in this particular type of recording session. In this way, we can not say that adding action can't help the actual learning of L2 speech. What action did conclusively *not* do in this study, was help the pronunciation spontaneously when the students acted. It is possible that the students were not aware of the action's role in the speech/text they were given, or the acting may have been random without really offloading the semantic demands onto their body. Also based on the interviews, all the participants often tell of the struggles with the suprasegmental phenomenon of tonemes, as opposed to the segmentals that are vowels and consonants. This could mean that their struggles with this particular part of Norwegian could come in the way of focusing, or practicing the phonemes they've learned, meaning their cognitive load could increase, though there is no conclusive evidence that this is the case. The importance of action on the accuracy of sounds might also be less impactful for such experienced learners, where improvement is slower than the starting point. The presence of action might give room to use speech in a more emotional state, but it seems it can also hinder focus of correct or precise pronunciation.

The accuracy rating of all the participants varied individually, where some had greater improvement of the action condition, and some had greater improvement in the non-action condition. Overall for all participants we saw that there was improvement over time for both conditions, especially in certain phonemes that the students expressed difficulties with, with the overall result of both conditions ending in the exact same percentage of accuracy in the final session. This tells us that the use of action does not negatively affect pronunciation accuracy. It might prove to be some challenge to add this type of action spontaneously during speech, and in a somewhat unusual recorded situation. Challenges are, in language learning, not necessarily bad, as it might create better understanding for the L2 speaker in how they pronounce and utilise their L2, and to get used to "being emotional" in their L2.

The lack of direct positive effect of action on the written speech also does not automatically entail a negative effect on accentedness and comprehensibility. There seems to be a slightly higher rating for non-action, although by very small measurements. The ongoing difference between action and non-action is present enough that it is possible to claim that adding action to a spontaneous situation (and with written, instructed speech) might tax the cognitive load somewhat in speech. Raters had some favouring towards non-action. The reasons could be many, though it is possible that the use of action had the participants sacrifice some of the pronunciation focus, and rather put more cognitive energy into the action and emotion that followed.

The participants also expressed the usefulness of speaking their L2 in everyday, private life, along with instructional learning, getting corrected, and importantly, allowing themselves to fall into the "Norwegianness" of the language, exaggerating melody and pronunciation of different elements, to teach the body a new normal, what some of the students referred to as "body memory". All of these reported and observed factors seem to be of aid in the improvement of sounds, accentedness and comprehensibility over time. The strongest finds that suggest this, is the improvement of the final R, and the general improvement for all participants on accuracy rates, and in accentedness and comprehensibility ratings. The action presented in this study is only one type of action, and in this case, it has shown that the spontaneous situation including action might not help pronunciation.

Finally, we have seen that action shows the same consistent results on the relationship between accentedness and comprehensibility. A noticeable accentedness score does not necessarily entail low comprehensibility. The sound accuracy shows a somewhat stronger improvement than accentedness scores overall. Interestingly, the numbers showing results of sound accuracy for the individual participants, and the numbers showing results of accentedness does not correspond. The sound accuracy percentage had a higher improvement overall than the accentedness and comprehensibility rating. The individual participants also had sound accuracy rates greater than accentedness and comprehensibility ratings. There is a sign of some consistency in numbers. One example being Margret, showing a consistent decrease in accentedness while gaining in sound accuracy, and then receiving higher accent score in the third session, as well as receiving less sound accuracy score

Table 14

Margret's accentedness rating versus sound accuracy

	Action			Non-action		
	S1	S2	S3	S1	S2	S3
ACC	4,1	3,75	4,6	4,6	4,4	4,9

Action			Non-Action		
S1	S2	S3	S1	S2	S3
62,2%	77,3%	69,8%	62,2%	75,4%	75,4%

We can see that the accentedness score, both in the action condition *and* the non-action condition, from S1 to S2 improves by going down, and sound accuracy goes up. Then, from S2 to S3 the accentedness score increases again, telling of a heavier accent, while the sound accuracy also decreases for the action condition. However, the accentedness score ends up being less favourable than initially, in both conditions, while the sound accuracy with the action condition ends up better than in S1, and remains stable from S2 to S3 in the non-action condition.

Interestingly, we can see similar results happening in comprehensibility, noticing how in the non-action condition, there is some improvement from S1 to S2, which then remains stable from S2 to S3, just like for sound accuracy. Comprehensibility scores in the action condition also improves, before showing some decrease in score.

Table 15

Margret's comprehensibility rating

	Action			Non-action		
	S1	S2	S3	S1	S2	S3
COMP	7,1	8	7,45	7,6	7,9	7,9

This particular example shows that there is a related relationship between sound accuracy and accentedness and comprehensibility. Although not perfectly matching, the results do show that there is a tendency for raters to be aware of higher sound accuracy in an individual. To add to the evidence, we can see the same tendency in Alexandrine.

Table 16

Alexandrine's accentedness and comprehensibility rating versus sound accuracy

	Action			Non-action		
	S1	S2	S3	S1	S2	S3
ACC	4,9	5,1	4,3	4,8	4,6	4,1
COMP	7,4	6,7	7,6	7,15	7,19	7,9

Action			Non-Action		
S1	S2	S3	S1	S2	S3
64,1%	64,1%	75,4%	69,8%	71,6%	81,1%

The results here show some consistency between the two measurements of pronunciation. In the non-action condition, sound accuracy improves somewhat steadily over time, with a higher jump in accuracy from S2 to S3. We can see the same for comprehensibility and for accentedness

ratings, with a rating showing favourable, steady results, with a larger difference between S2 and S3. In the action condition, the accentedness and comprehensibility rating follows one another in the first increase of ACC score and decrease in COMP score from S1 to S2, while the sound accuracy stayed the same from S1 to S2. The numbers still end up with a more favourable score in all measurements in S3.

We can see signs of this consistency with Thyra as well, though the consistency is somewhat lower. Accentedness and comprehensibility scores show some different tendencies in progression in the action condition. Though if we compare accuracy rate and comprehensibility score, we can see a steady improvement over time. With the non-action condition, accentedness and comprehensibility scores show first a disfavourable development, before improving in accentedness, but still decreasing in comprehensibility. Meanwhile, sound accuracy rates first show a favourable increase, before decreasing somewhat.

Table 16

Thyra's accentedness and comprehensibility rating versus sound accuracy

	Action			Non-action		
	S1	S2	S3	S1	S2	S3
ACC	4,9	4,9	4,5	4,2	4,04	4,1
COMP	7,1	7,3	7,6	8,2	8	7,67

Action			Non-Action		
S1	S2	S3	S1	S2	S3
56,6%	64,1%	79,2%	60,3%	71,6%	67,9%

What we can take from these numbers is that there is a certain pattern to the relationship between sound accuracy and the rating scores of accentedness and comprehensibility, even though the

relationship is not one to one. The results also show that, with the inclusion of action, there is no steady, consistent improvement, but improvement is present.

With these findings I can neither conclude nor dismiss that the use of action leads to improvement on the pronunciation accuracy of nonnative sounds, and the overall oral proficiency measured by comprehensibility and accentedness. The numbers tell us, however, that action does not aid pronunciation in the moment, and possibly overcomplicates the processing for the learners. The role of action nevertheless proves to be of a certain aid in many studies, and seems to improve the engagement and enjoyment of learning a language. The study of action is one but many parts of the theory of EC, an expanding field which proves to be of great interest, and applicable to a variety of fields. I will get further into this in the next section.

5.4 Remarks on Embodied Cognition and action, remarks on teaching Norwegian

Through this experiment, I have attempted to contribute to the field within EC and language learning. The experiment was narrowed down to a very specific form of EC, through the use of "action" during speech, tied with the discipline of acting, in an attempt to explore the possible effects on sound accuracy, accentedness and comprehensibility.

There are many EC studies involving various types of action tied with language learning, showing the various benefits of incorporating the body when learning something not intrinsically physical. The type of action in this study can be described as a quite "subtle" kind of action, with the participants sitting down, reading from a sheet, only relying on what their abilities to spontaneously be "directed" into a certain emotion, reaction, or intention. It is worth asking if this would be considered enough of an "action"? Often in these cases, what little physical, visible action they did would be limited to small nods, sighs, frowns, as well as raising or lowering the voice. I would argue that what the instructions to add action did for the participants was to trigger an active engagement into the text and the meaning, which is an essential and integral to acting as a craft. What this type of action within acting also can do, which echoes the findings in this study, is to claim a lot of concentration from an actor. There lies work behind the process of the mind (and body) of an actor when on stage. This project did not involve a stage, and attempted to find a minimal effort for the acting students, not requiring any research for a role or play, but keep it to the bare bones of putting the imagination going. Still, the results

implied that this type of action possibly taxed their cognitive processing when reading and speaking the text.

In hindsight, there are several possibilities in how action in the form of bodily movements could have been incorporated, and these have been considered. There could have been other effects on the speech if including bodily movements like standing up while uttering the sentences, having an actor partner (either the researcher, or another person not being studied) to speak towards, or do a closer replication of the use of practice sentences in the classes that I observed. There could even be some integration of gesture, alongside the use of acting (though this potentially could be too cognitively taxing as well, and difficult to study two simultaneous actions' effect on pronunciation). Earlier having shown that including gesture, body language and proprioception in speech of L2 (Bach-Marquès & Carrera-Sabaté, 2019), a more direct use of the body could be of further aid. However, it is worth asking if making the smallest differences in pronunciation would be visible through gesture, considering the students' intermediateness. What kind of body language or bodily movements could help pinpointing the smallest changes and diversions from native speech? These questions are not immediately obvious, and might be protested by Flege and others in their claims that learning speech identically to native speech could be impossible (Flege, 2021; Derwing & Munro, 2009; Galante & Thomson, 2017). At the very least, considering the participants' frustrations with tonemes in Norwegian, there could be promising results in the future using body language and gesture resembling the drawing intonation contour and tones, as suggested by some research (Bach-Marquès & Carrera-Sabaté, 2019).

There are also other factors not related to the use of action that could have an impact on their progression. The reports from the participants of rapid learning of Norwegian in the beginning of their language learning which then stagnated, leading to progression being slower and slower correspond with the findings, in that the progression is present, though relatively minor. They also expressed the difficulties of keeping improving in their pronunciation when their comprehensibility was high in that they were always being understood by their peers. This could easily lead to what they called "laziness", or that possibilities of improvement came more seldom, they were less corrected by others, and their errors were very subtle and hard to identify and practice. Action in the form of acting does seem to have certain benefits for speech and L2 use in some ways. Considering Margret's thoughts on the use of Norwegian on stage, like allowing herself to not be too hung up on the exact pronunciation, and "go for it", this could imply that the use of an L2 in a stage performance, or adding action to the L2 speech, could be

an additional cognitive load for the L2 learner, in terms of precision in speech, while at the same time providing a certain "flow" in speech, which could make them more accented, but also more fluent (see various definitions of fluency in Galante & Thomson, 2017). This way, action itself does not necessarily aid in speech sound accuracy, but could provide with a motivation during speech, to use emotion and expressions, which in turn have been proven to improve pronunciation (Estrada, 2004, 2007).

Finally, though there are individual and more global challenges to learning Norwegian for all learners (as there is with learning any language), there might be unique challenges for learners who speak a very familiar language that is intelligible. One of the participants remarked that she had at times been using Danish words and "Norwegified" them, in thinking both languages utilised the same word in the same way, which is not always the case. It is possible that this does not only go for vocabulary, but also for pronunciation, that Danish sounds only get approximately closer to the Norwegian equivalent, not knowing exactly when one has "hit the mark". This, however, could easily be said for other learners as well, who often have to radically change their vocabulary and pronunciation habits.

To learn Norwegian as a Danish speaking person in itself is not hard compared to other learners, considering their great similarities. Still, there will always be difficulties. And considering the improvement of these particular students over the course of their studies, to completely rule out the role of action and the importance of EC in such a discipline as acting should not be done. There are still possibilities to be done in future research, which I will briefly suggest below.

5.5 Future research

As mentioned, there are several possibilities in incorporating various types of gestures, body language and other physical movements in this type of experiment, in an attempt to further engage the participants in their actions. It might be beneficial to include an exercise that the participants already know well, either from their language classes, or other classes that can easily incorporate the use of speech and short bursts of emotion and action. Considering the promising results mentioned in this thesis from earlier research on the use of many kinds of bodily movements, this might prove to give an even better result. This can surely be tried out with Norwegian tonemes as well, considering the difficulties learning them.

This study included very few participants, which makes it hard to apply the results from this experiment onto a larger context. It would be of interest to do such a study in a larger context, and possibly over a longer span of time. To study these kinds of acting students from the very start of their language learning to the end of the last class would certainly give a larger body of data.

I also want to briefly discuss the issues of clearly defining "acting" as an artistic practice. To do "acting", or "to act", will naturally have a diverse range of interpretation and personal styles for actors, where all actors may have small or big differences in interpretation of how to act a certain situation or feeling (think how Nicolas Cage would approach a scene about betrayal vs. how Humphrey Bogart would have done it). This, in turn, will then affect how much acting will trigger factors such as emotions and different kinds of visible action. A setting of asking an actor, while sitting down and reading from a paper might put some limitation on the range of acting, as well as action. It could also be worth considering for future discussion the difference between acting and engaged conversations in real life, and how this may affect pronunciation training. In both cases, a certain type of emotional reaction, which would entail action, would be present. The specifics on how this could be studied are not of my knowledge as of now, but it is nonetheless a suggestion of mine.

Considering Estrada's (2004, 2007) findings, a narrower instruction of how to do action, or convey a certain emotion, could possibly have yielded different results. However, considering how action in this experiment showed improvement alongside non-action, this suggestion does not imply the definition and use of acting and action in this thesis as a failure, but rather being a suggestion to continue research involving emotions, actions, and EC.

5.6 Limitations

Section **5.6** will address some of the challenges along the way and discuss the limitations of the research design. These could hopefully be used to improve the project, or further encourage further study of a similar topic, if it is to be repeated in any form.

As mentioned, this study only contains a few participants, which means the results can not automatically be assumed to be transferable to other studies or contexts. There was also no

control group in this study, mostly due to the nature of the educational set up, which didn't include any Danish students that only practiced pronunciation without any particular action/acting aspect. This makes the effect on action/acting on the participants' learning development more uncertain.

If done again, I would change the sentence sheets with and without action given to the participants, randomizing the sentences each time, which was not done this time. There is also the possibility that the word "action" (NOR: *handling*) in the questionnaire would benefit from either a clearer explanation from me as a researcher, or could have been changed to cater to these particular students' previous knowledge of terms and experiences.

Another thing to note, it is difficult to navigate with an academic study such as this in a world including artistic fields such as acting. It requires a knowledge of both sides of the study, which I luckily had, in that one can understand the nuances that arise in a study involving the creative and artistic (naturally, this does not mean language studies are not creative or artistic). This, however, is not necessarily a negative aspect of such a study. In fact, it only means there is a lot more potential to be further studied and discovered, with many trials and errors on the way, on the scientific journey to combine several disciplines of the academic world.

Finally, I did not have enough time to add more research and background on the field of acting, which I believe would benefit this thesis, especially for those who do not know much about theatre and acting. If the readers of this thesis might find the theme interesting themselves, I suggest further reading in Kaplan (2007), Krashner (2000), Longwell & Meisner (1987) or Holter et al. (2017), as potential reading on certain acting methods and what they entail. For future research, I hopefully will be able to include more information on this.

6. Conclusion

This study has explored the use of action as a tool for language learning, where action was realized through the use of acting. The study followed three young adult Danish acting students learning Norwegian, and who had been learning the language since their first semester in school. At the point when the study commenced, the students were in their second year and had reached a certain proficiency already.

I have presented the theory of Embodied Cognition and research that explores and explains what the theory entails, and its potential within linguistics, as well as several fields of study. A multitude of studies has shown a favourable effect of EC on many linguistic topics, particularly language learning. It is apparent that the use of the body is central to not only learning, but processing of information overall. Researchers have argued greatly for the recognition of EC, to see mind and body as our cognition altogether, that thought and cognition is a physical process (Lakoff, 2012), and that neural and non-neural processes have a dynamic interplay (Foglia & Wilson, 2013), tying EC to better learning and education (Ionescu & Vasc, 2014), showing that body and mind work together in, among other things, learning new languages, and arguing that as a complex and multimodal field of study it should be studied in equal terms, and focus on learning as something relational to people and the environment (Atkinson, 2010). EC is supported and motivated by observations around gesture use, visual perception and bodily movement, mirror neurons, bodily movements in cognitive tasks etc. Even the use of metaphors echoes our cognitive processing of the world, as we see, hear, feel, taste and experience it (Foglia & Wilson, 2013; Lakoff & Johnson, 1980). Even within artistic sciences and disciplines like acting, we find engagement within EC, creating terms like the "bodymind", an embodied practice where one engages both the physical and the cognitive, tying it with experience and awareness (Zarrilli, 2004).

I have argued in this experiment that, based on the theory of EC, the body participates in the L2 pronunciation online, meaning that producing speech produces movement. Therefore, it partly offloads the semantic information onto the bodily movements, which opens the possibility for lightening the cognitive load, making it so that the participants could have more cognitive resources to focus on pronunciation itself. The goal of the study has been to explore and add further knowledge in how acting can be used and incorporated with language learning and linguistics as a field. It also attempted to find evidence for possibilities in action lightening the cognitive load. To explore these questions, the thesis addressed two research questions, namely: 1) Is there a substantial or noticeable effect of the use of action vs. non-action on the pronunciation accuracy of nonnative sounds? and 2) Is there a substantial or noticeable improvement with the use of action vs. non-action with overall oral proficiency on the measures of comprehensibility and accentedness?

To answer these questions, I collected speech data, answers from questionnaires, and did interviews from the three Danish participants, as well as collecting native Norwegian speech. The last part of the data collection consisted of a rating session on accentedness and comprehensibility from three native Norwegians unfamiliar to all the speech participants.

The results from this study agree with some of the previous research presented. Studies have shown that the use of drama and role plays in the classroom are reported to be enjoyable and fun for students, who also report feeling more confident in their speaking (Abenoja & DeCoursey, 2019), and increase oral fluency (Galante & Thomson, 2017). However, neither of the studies found any significant effect on accentedness or comprehensibility. The improvement of the participants in this study was quite minor. This is not surprising considering the L2 learners were already quite proficient, and L2 sound categories are formed slowly (Flege, 2021).

The results found in this experiment showed that there was improvement for all participants over time in terms of both accentedness and comprehensibility. Speech accuracy had mixed results in their improvement, often rooted in the fact that many of the phonemes in focus were pronounced with high accuracy from the first session. In particular, the final R showed promising improvement over time for all participants. This was also the single phoneme for all participants which they struggled the most with, being the lowest in accuracy rate of all the phonemes in the first session. For Thyra and Alexandrine, the increase in accuracy for the final R was present for every session, while for Margret, there was an increase in accuracy from session one to session two, but then decreased in session three, but only in the action condition, while the non-action condition increased from session one to session two, and remained the same accuracy in the third session. Interestingly, this difference between the participants was similar in the accentedness and comprehensibility rate, with the same increase and decrease pattern for Margret, and only in the action condition. This does not mean that Margret in particular lost any pronunciation progression, as shown by the non-action conditions across the data points. It might just mean that the action condition on this day were not working in her favour, even less so than the previous session.

Results showing individual phoneme accuracy rate, and the individual participant's improvement on accentedness and comprehensibility did not show a clean, steady improvement, and showed both increase and decrease in all variables. However, the final session showed a higher accuracy rate, lower accentedness and higher comprehensibility in most of the cases, and overall accuracy

rate and accentedness and comprehensibility rates showed favourable results. This shows that the use of action did yield some favourable results in their pronunciation training. However, with non-action showing equally, and often more favourable results, the use of action was not the *cause* of the improvement.

Considering the questionnaire and the interviews, the participants found the use of action a favourable and helpful tool in improving their speech and pronunciation. This was along with the mere daily use of the language, especially in personal situations, which seemed to be a high motivator to continue in using the language, and further sparked curiosity in improving pronunciation. When reporting difficulties, tonemes were often mentioned, as well as the issue of slowing down, often due to their Norwegian being highly comprehensible, and "fluent enough" to native Norwegians. This could potentially lead to being less corrected when needed, and the L2 speakers being less aware of their minor deviations from L1 speech.

The data did not show any conclusive proof that the use of action resulted in lightening the cognitive load, as shown that the action condition rarely yielded more favourable results than the non-action condition. In fact, quite often the non-action condition showed more favourable results, across all data measurements, although not substantially higher. This implies that during these recording sessions, the inclusion of action may have taxed the cognitive load, rather than lightening it. This shows that the use of action did not have a favourable effect over non-action on the accuracy of pronouncing nonnative sounds. Considering the results on accentedness and comprehensibility rating, the numbers imply clear improvement with overall proficiency, but action had no advantage over non-action. These answers do not automatically mean that action has no effect on learning L2 speech, but it does not help pronunciation spontaneously while acting. I will then argue that, although not all the findings could be satisfyingly accounted for, there are still possibilities tied with the use of action, and thereby acting, in the study on pronunciation improvement.

This thesis has had the goal of contributing to the field of language learning, Embodied Cognition, and promoting the incorporation of theatre and acting in these fields of study. Embodied Cognition as a theory leaves still a lot to be explored, and shows promising results in various studies focusing on our learning. As a researcher, I have found the notion of incorporating the body with cognitive learning appealing and highly interesting. Finally, the use of action as a definition can be realized in many different ways, which provides the opportunity

for many kinds of studies involving action. When faced with difficulties in learning language, incorporating embodied practice can facilitate the learning and processing, and if it doesn't directly help factors like pronunciation in nonnative sounds, it at the very least makes the learning more enjoyable.

References

Abenoja, Z. M. K. M., & DeCoursey, M. (2019). Using drama activities to teach beginner's French to Chinese students at a tertiary institution in Hong Kong: An exploratory case study. *Studies in Second Language Learning and Teaching*, *9*(4), 711–736. https://doi.org/10.14746/ssllt.2019.9.4.7

Acton, W., Baker, A., Burri, M., & Teaman, B. (2013). Preliminaries to haptic-integrated pronunciation instruction. In J. Levis & K. LeVelle (Eds.), *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference,* August 2012. (pp. 234–244). Ames, IA: Iowa State Universit

Andersson, E. (1994). Swedish. In E. König & J. van der Auwera (Eds.) *The Germanic Languages* (271-312). Routledge
Askedal, J. O. (1994). Norwegian. In E. König & J. van der Auwera (Eds.) *The Germanic Languages* (219-270). Routledge

Atkinson, D. (2010). Extended, Embodied Cognition and Second Language Acquisition. *Applied Linguistics,* 31-5, p. 599–622. doi:10.1093/applin/amq009

Bach-Marquès, J. & Carrera-Sabaté, J. (2019). Gestures and prosody to enhance the pronunciation of Catalan vowels by English native speakers learning Catalan. *International Workshop Embodied Speech*. Paris, France.

Baills, F. & Prieto, P. (2021). Embodying rhythmic properties of a foreign language through hand- clapping helps children to better pronounce words. *Language Teaching Research.* 10.1177/1362168820986716.

Baills, F., Suárez-González, N., González-Fuente, S., & Prieto, P. (2019). Observing and producing pitch gestures facilitates the learning of mandarin chinese tones and words. *Studies in Second Language Acquisition, 41*(1), 33-58. doi:10.1017/S0272263118000074

Barsalou, L.W. (2003). Situated simulation in the human conceptual system. *Language and Cognitive Processes*, 18 (5-6), 513–562. https://doi.org/10.1080/01690960344000026

Basbøll, Hans (2005). *The Phonology of Danish.* New York: Oxford University Press.

Bickhard, M. H. (2008). Is Embodiment Necessary? In P. Calvo & A. Gomila (Eds.), Handbook of Cognitive Science: An Embodied Approach, San Diego, Elsevier. https://doi.org/10.1016/B978-0-08-046616-3.00002-5.

Biggs, J. (1996). Western misperceptions of the Confucian-heritage learning culture. In D. Watkins & J. Biggs (Eds.), *The Chinese learner: Cultural, psychological and contextual influences* (pp. 45-68). Hong Kong: Comparative Education Research Centre and Australian Council for Educational Research

Boersma, P. & Weenink, D. (2023). Praat: doing phonetics by computer [Computer program]. Version 6.2.23, retrieved 27 March 2023 from http://www.praat.org/

Bonda, E., M. Petrides, S. Frey, & A. Evans. (1994). Frontal cortex involvement in organized sequences of hand movements: Evidence from a positron emission topography study. *Society for Neurosciences Abstracts.* 20: 353.

Borghi, A. M., Glenberg, A. M., & Kaschak, M. P. (2004). Putting words in perspective. *Memory & Cognition, 32*(6), 863–873. https://doi.org/10.3758/BF03196865

Burri, M. & Baker, A. (2016). Teaching rhythm and rhythm grouping: The butterfly technique. *English Australia Journal: the Australian journal of English language teaching,* 31 (2), 72-77.

Burri, M., Baker, A. (2019) "I never imagined" pronunciation as "such an interesting thing": Student teacher perception of innovative practices. *Int J Appl Linguist*. 2019; 29: 95– 108. https://doi.org/10.1111/ijal.12247

Carless, D. (2011). *From testing to productive student learning: Implementing formative assessment in Confucian-heritage settings*. New York: Routledge

Clark, A. (2001). Mindware: An Introduction to the Philosophy of Cognitive Science. Oxford University Press.

Cook, V. (1999). Going beyond the native-speaker in language teaching. TESOL Quarterly 33 (2): 185–209

Crystal, D. (2003). English as a Global Language (2nd ed.). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511486999

De Marco, A. (2020). Teaching the Prosody of Emotive Communication in a Second Language. In C. Savvidou (Red.), *Second Language Acquisition - Pedagogies, Practices and Perspectives* (p. 5-22). IntechOpen. doi: 10.5772/intechopen.77639

DeKeyser, R., & Larson-Hall, J. (2005). What does the critical period really mean? In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 88–108). New York: Oxford University Press.

Delsing, L. O., & Lundin-Åkesson, K. (2005). Håller språket ihop i Norden? : En forskningsrapport om ungdomars förståelse av danska, svenska och norska. https://doi.org/10.6027/TN2005-573

Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching,* 42, 476–490. doi:10.1017/

Elvira-García W. (2018). Extracts f0 from points and matching interval label [praat script]. Retrieved from https://www.ub.edu/phoneticslaboratory/praatscripts/extract_f0_from_points__and_matching_intervalLabel.praat.

Esteve-Gibert, N., Del Mar Suárez, M., Vasylets, O., & Serrano, R. (2019). Children's use of tactile input when acquiring non-native phonological contrasts. *Presentation at XIV International Symposium of Psycholinguistics.* DOI: 10.13140/RG.2.2.30102.60482/1

Estrada Medina, M. (2004). L'expression de l'émotion et la correction phonétique. l'exemple de la surprise. In *Le français face aux défis actuels. Histoire, Langue et Culture* (1 ed., pp. 319-329).

Estrada, M. (2007). Incidence de la prosodie sur la structuration de la matière phonique: l'exemple de la surprise. *Deuxième Colloque International de Didactique Cognitive des Langues*, en Toulouse, France. Volume: 1

Farrell, T. S., & Martin, S. (2009). To teach standard English or world Englishes?3 A balanced approach to instruction. In *English Teaching Forum*, Vol. 47, No. 2. Washington, DC: US Department of State. Bureau of Educational and Cultural Affairs, Office of English Language Programs

Ferrari, P. F., Coudé, G. (2018) Mirror Neurons, Embodied Emotions, and Empathy. In K. Z. Meyza & E. Knapska (Eds.), Neuronal Correlates of Empathy, 67-77. Academic Press. https://doi.org/10.1016/B978-0-12-805397-3.00006-1.

Flege, E. F., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *J Acoust Soc Am,* 1 May 1995; 97 (5). 3125–3134. https://doi.org/10.1121/1.413041

Flege, J. (1992b). The intelligibility of English vowels spoken by British and Dutch talkers. DOI: 10.1075/sspcl.1.06fle

Flege, J. (1995). Second language speech learning: Theory, findings and problems. (229-273). York Press

Flege, J., & Bohn, O. (2021). The Revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress* (pp. 3-83). Cambridge: Cambridge University Press. doi:10.1017/9781108886901.002

Foglia, L. & Wilson, R. (2013). Embodied Cognition. *WIREs Cogn Sci* 2013, 4:319–325. doi: 10.1002/wcs.1226

Galante, A. & Thomson, R.I. (2017), The Effectiveness of Drama as an Instructional Approach for the Development of Second Language Oral Fluency, Comprehensibility, and Accentedness. *TESOL Q, Volume 51, issue 1*, 115-142. https://doi.org/10.1002/tesq.290

Gilbert, J. B. (2008). *Teaching Pronunciation: Using the Prosody Pyramid.* Cambridge University Press

Giles, H. (1970) Evaluative reactions to accents, *Educational Review*, 22:3, 211-227, DOI: 10.1080/0013191700220301

Glenberg, A. M., Sato, M., Cattaneo, L., Riggio, L., Palumbo, D., & Buccino, G. (2008). Processing abstract language modulates motor system activity. *The Quarterly Journal of Experimental Psychology, 61*(6), 905–919. https://doi.org/10.1080/17470210701625550

Gluhareva, D., & Prieto, P. (2017). Training with rhythmic beat gestures benefits L2 pronunciation in discourse-demanding situations. *Language Teaching Research*, 21(5), 609–631. https://doi.org/10.1177/1362168816651463

Grønnum, N. (2008). Hvad er det særlige ved dansk som gør det svært at udtale og forstå for andre? - Første del: enkeltlydene. *Mål og Mæle*, 31-1, p. 15-20.

Haberland, H. (1994). Danish. In E. König & J. van der Auwera (Eds.), *The Germanic Languages* (313-348). Routledge

Hardison, D. (2003). Acquisition of second-language speech: Effects of visual cues, context, and talker variability. *Applied Psycholinguistics, 24*(4), 495-522. doi:10.1017/S0142716403000250

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). 'Somatotopic representation of action words in human motor and premotor cortex,' Neuron 41: 301–7.

Hirata, Y., Kelly, S. D. (2010). Effects of Lips and Hands on Auditory Learning of Second-Language Speech Sounds. *Journal of Speech, Language, and Hearing Research,* 53, p. 298–310

Holter, L. A., Strømdahl, T., Stene, Ø., Lid, T. V., & Gunnes, T. (2017) Skuespillerens arbeid med tekst. Kunsthødskolen i Oslo http://hdl.handle.net/11250/2454705

Hulse, B., & Owens, A. (2019) Process drama as a tool for teaching modern languages: supporting the development of creativity and innovation in early professional practice. *Innovation in Language Learning and Teaching,* 13:1, 17-30, DOI: 10.1080/17501229.2017.1281928

Iizuka, T., Nakatsukasa, K. and Braver, A. (2020), The Efficacy of Gesture on Second Language Pronunciation: An Exploratory Study of Handclapping as a Classroom Instructional Tool. *Language Learning,* 70: 1054-1090. https://doi.org/10.1111/lang.12415

Ionescu, T., & Vasc, D. (2014). Embodied Cognition: Challenges for Psychology and Education. Procedia - Social and Behavioral Sciences, Volume 128, p. 275-280, doi.org/10.1016/j.sbspro.2014.03.156.

Jia, G., Strange, W., Wu, Y., Collado, J., & Guan, Q. (2006). Perception and production of English vowels by Mandarin speakers: Age related differences vary with amount of exposure. *Journal of the Acoustical Society of America*, 119(2), 1118–1130.

Kaplan, D. (2007) The Five Approaches to Acting Series: The Collected Series. Hansen Publishing Group

Kartushina, N., Hervais-Adelman, A., Frauenfelder, U. H., and Golestani, N. (2015) The effect of phonetic production training with visual feedback on the perception and production of foreign speech sounds. *The Journal of the Acoustical Society of America* 138, 817-832. https://doi.org/10.1121/1.4926561

Kelly, S. D. & Hirata, Y. (2017) What neural measures reveal about foreign language learning of Japanese vowel length contrasts with hand gestures. In S. Tanaka et al. (eds.), *New Development in Phonology Research: Festschrift in Honor of Haruo Kubozono* (音韻研究の新展開:窪薗晴夫教授還暦記念論文集), pp. 278-294. Tokyo: Kaitakusha.

Kelly, S. D., Hirata, Y., Manansala, M., Huang J. (2014). Exploring the role of hand gestures in learning novel phoneme contrasts and vocabulary in a second language. *Frontiers in Psychology*, 5. DOI 10.3389/fpsyg.2014.00673

Kiefer, M., & Trumpp, N.M. (2012). Embodiment theory and education: The foundations of cognition in perception and action. *Trends in Neuroscience and Education, 1*, 15-20. https://doi.org/10.1016/j.tine.2012.07.002

Krasner, D. (2000). Strasberg, Adler and Meisner: Method Acting. In Hodge, Alison (Eds). Twentieth Century Actor Training. London and New York: Routledge. pp. 129–150.

Kristoffersen, G. (2000). The Phonology of Norwegian, Oxford University Press, p. 11-19

Kushch, O. (2018). Beat gestures and prosodic prominence: impact on learning. *TESI DOCTORAL UPF*

Lakoff, G. (2012), Explaining Embodied Cognition Results. Topics in Cognitive Science, 4: 773-785. https://doi.org/10.1111/j.1756-8765.2012.01222.x
Lakoff, G., Johnson, M. (1980) *Metaphors We Live By*. Chicago: University of Chicago Press

Lambert, W.E., Hodgson, R.C., Gardner, R.C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *Journal of abnormal and social psychology, 60*, 44-51. . https://doi.org/10.1037/h0044430

Lee, B., Plonsky, L., Saito, K. (2020) The effects of perception- vs. production-based pronunciation instruction. *System*, 88. https://doi.org/10.1016/j.system.2019.102185.

Lenneberg, E.H. (1967). *Biological foundations of language.* New York: Wiley.

Li, P., Baills, F., & Prieto, P. (2020). Observing and producing durational hand gestures facilitates the pronunciation of novel vowel-length contrasts. *Studies in Second Language Acquisition, 42*(5), 1015-1039. doi:10.1017/S0272263120000054

Llompart, M., & Reinisch, E. (2019). Imitation in a Second Language Relies on Phonological Categories but Does Not Reflect the Productive Usage of Difficult Sound Contrasts. *Language and Speech,* 62(3), 594–622. https://doi.org/10.1177/0023830918803978

Longwell, D., & Meisner, S. (1987). Sanford Meisner on Acting. New York: Random House.

Lord, G. (2008). Podcasting Communities and Second Language Pronunciation. *Foreign Language Annals.* 41. 364 - 379. 10.1111/j.1944-9720.2008.tb03297.x.

Lord, G.. (2005). (How) Can We Teach Foreign Language Pronunciation? On the Effects of a Spanish Phonetics Course. *Hispania* 88.3: 557-67

Merlin, B. (2007) *The Complete Stanislavsky Toolkit.* Quite Specific Media Group Ltd

Morett, L. M., Chang, L. Y. (2015) Emphasising sound and meaning: pitch gestures enhance Mandarin lexical tone acquisition. *Language, Cognition and Neuroscience*, 30:3, 347-353, DOI: 10.1080/23273798.2014.923105

Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45, 73–97. doi:10.1111/0023-8333.49.s1.8

Munro, M. J., & Derwing, T. M. (2015). A prospectus for pronunciation research in the 21st century. *Journal of Second Language Pronunciation, Volume 1, Issue 1*, Jan 2015, p. 11 - 42. https://doi.org/10.1075/jslp.1.1.01mun

Munro, M. J., & Derwing, T. M. (2020). Foreign accent, comprehensibility, and intelligibility, redux. *Journal of Second Language Pronunciation, Volume 6, Issue 3, Nov 2020,* p. 283 - 309. https://doi.org/10.1075/jslp.20038.mun

NHI (Norges Helseinformatikk). (16.02.2023). *Asperger Syndrom.* https://nhi.no/. https://nhi.no/sykdommer/barn/autisme/asperger-syndrom/?page=3

Oberman, L. M., & Ramachandran, V. S. (2007). The simulating social mind: The role of the mirror neuron system and simulation in the social and communicative deficits of autism spectrum disorders. *Psychological Bulletin, 133*(2), 310–327. https://doi.org/10.1037/0033-2909.133.2.310

Offerman, H. M., Olson, D. J. (2016). Visual feedback and second language segmental production: The generalizability of pronunciation gains. *System*, Volume 59, 2016, P. 45-60. https://doi.org/10.1016/j.system.2016.03.003.

Olson, D. (2014). Benefits of visual feedback on segmental production in the L2 classroom. *Language Learning and Technology.* 18. 173-192.

Ozakin, A. S., Xi, X., Li, P. & Prieto, P. (2022) Thanks or Tanks: Training with Tactile Cues Improves Learners' Accuracy of English. I*nterdental Consonants in an Oral Reading Task, Language Learning and Development,* DOI: 10.1080/15475441.2022.2107522

Patel-Grosz, P., Grosz, P. G.; Kelkar, T. & Jensenius, A. R. (2018). Coreference and disjoint reference in the semantics of narrative dance. In Sauerland, Uli & Solt, Stephanie (Red.), *Proceedings of Sinn und Bedeutung 22, vol. 2, ZASPiL 61*. Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS). s. 199–216.

Patel-Grosz, P., Grosz, P. G.; Kelkar, T. & Jensenius, A. R. (2022). Steps towards a semantics of dance. *Journal of Semantics*. ISSN 0167-5133. 39(4), s. 693–748. doi: 10.1093/jos/ffac009.
Penfield, W. (1965). Conditioning the uncommitted cortex for language learning, *Brain*, Volume 88, Issue 4, November 1965, Pages 787–798, https://doi.org/10.1093/brain/88.4.787

Pylyshyn, Z. W. (1980). Computation and cognition: Issues in the foundations of cognitive science. *Behavioral and Brain Sciences, 3*(1), 111–169. https://doi.org/10.1017/S0140525X00002053

Rieser, J., A. Garing, and M. Young. (1994). 'Imagery, action and young children's spatial orientation: It's not being there that counts, it's what one has in mind,' Child Development 45: 1043–56.

Rizzolatti, G. & Craighero, L. (2004). The Mirror-Neuron System. *Annual review of neuroscience*. 27. 169-92. 10.1146/annurev.neuro.27.070203.144230.

Rizzolatti, G., Fogassi, L. & Gallese, V. (2001) Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat Rev Neurosci* **2**, 661–670. https://doi.org/10.1038/35090060

Saito, K. & Lyster, R. (2012), Effects of Form-Focused Instruction and Corrective Feedback on L2 Pronunciation Development of /ɹ/ by Japanese Learners of English. *Language Learning*, 62: 595-633. Language Learning Research Club, University of Michigan. https://doi.org/10.1111/j.1467-9922.2011.00639.x

Saito, K. (2011). Examining the role of explicit phonetic instruction in native-like and comprehensible pronunciation development: an instructed SLA approach to L2 phonology. Language Awareness. *Language Awareness,* Vol. 20, No. 1, p. 45–59. Taylor & Francis. DOI: 10.1080/09658416.2010.540326

Saito, K., & Wu, X. (2014). Communicative Focus On Form And Second Language Suprasegmental Learning: Teaching Cantonese Learners to Perceive Mandarin Tones. *Studies in Second Language Acquisition, 36*(4), 647-680. doi:10.1017/S0272263114000114

Saito, Y., & Saito, K. (2017). Differential effects of instruction on the development of second language comprehensibility, word stress, rhythm, and intonation: The case of inexperienced Japanese EFL learners. *Language Teaching Research,* 21(5), 589–608. https://doi.org/10.1177/1362168816643111

Savvidou, C. (2020). Second Language Acquisition - Pedagogies, Practices and Perspectives. IntechOpen. doi: 10.5772/intechopen.77639

Scherer, K.R., Banse, R., & Wallbott, H.G. (2001). Emotion Inferences from Vocal Expression Correlate Across Languages and Cultures. *Journal of Cross-Cultural Psychology, 32*, p. 76-92. DOI:10.1177/0022022101032001009

Shapiro, L. & Stolz, S. A. (2018). Embodied cognition and its significance for education. Theory and Research in Education, 1-21. https://doi.org/10.1177/1477878518822149

Spada, N. & Lightbrown, P. M. (2008) Form-Focused Instruction: Isolated or Integrated? *TESOL Quarterly* 42, no. 2 (2008): 181-207. Teachers of English to Speakers of Other Languages, Inc.

Sternberg, E. (2016) *NeuroLogic: The Brain's Hidden Rationale Behind Our Irrational Behavior.* Pantheon. Online excerpt available: https://www.psychologytoday.com/us/blog/neurologic/201601/why-is-yawning-contagious

Summerfield, Q. (1979). Use of Visual Information for Phonetic Perception. *Phonetica, 36:* 314-331

Teaman, B. D., & Acton, W. R. (2013). Haptic (movement and touch for better) pronunciation. In N. Sonda & A. Krause (Eds.), *JALT2012 Conference Proceedings*. Tokyo: JALT

Torp, A (2005). Nordiske sprog i fortid og nutid. I I. S. Sletten (red.). *Nordens sprog - med rødder og fødder* (s. 19-74). Nordisk Ministerråd. https://doi.org/10.6027/Nord2004-010

van Maastricht, L:, Hoetjes, M., van Drie, E. (2019). Do gestures during training facilitate L2 lexical stress acquisition by Dutch learners of Spanish? *International Conference on Auditory-Visual Speech Processing*. DOI: 10.21437/AVSP.2019-2

Wennersten, M. R. (2022). *Verdigrunnlag for språkovergang* [Informational document from art institution mentioned in thesis].

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review, 9*, 625-636. DOI:10.3758/BF03196322

Zarrilli, P. B. (2004). Towards a Phenomenological Model of the Actor's Embodied Modes of Experience. Theatre Journal , Dec., 2004, Vol. 56, No. 4, Theorizing the Performer (Dec., 2004), p. 653-666. Johns Hopkins University Press

Zhang, R., & Yuan, Z. (2020). Examining the effects of explicit pronunciation instruction on the development of L2 pronunciation. *Studies in Second Language Acquisition, 42*(4), 905-918. doi:10.1017/S0272263120000121

Zhang, Y., Baills, F. Prieto, P. (2018). Hand-clapping to the rhythm of newly learned words improves L2 pronunciation: Evidence from Catalan and Chinese learners of French. *Language Teaching Research.* 1-24. DOI: 10.1177/1362168818806531

Appendix A

Consent form (blank copy)

Are you interested in taking part in the research project

"Pronunciation in second language and second dialect: The case of Danish learning Norwegian"

Purpose of the project

You are invited to participate in a research project where the main purpose is to do research on language learning, specifically Danish learning Norwegian, in a theatre/acting institution.

This research project is a part of an MA thesis in linguistics, at the University of Oslo. The goal of the project is to investigate pronunciation of Norwegian among Danish learners. The focus of the project is the embodiment and character work of the actor, as a context for the active use and learning of Norwegian. The project contributes to knowledge about dialect and pronunciation learning.

Which institution is responsible for the research project?

Faculty of Linguistics at The University of Oslo is responsible for the project

Project leader and supervisor for the project is Peng Li and Haley De Korne at Multiling at The University of Oslo

Why are you being asked to participate?

You're asked to participate because you're a Danish student at a Norwegian acting school, and are currently having language classes in Norwegian. This is a unique situation where said language that is learned is very close to the first language/mother tongue, where learning of pronunciation is the main focus, rather than learning a new grammar.

What does participation involve for you?

For the participants, this involves being observed in language classes, as well as to be recorded (audio only) three times, over a period of time (3 months). The recording will involve reading

some example sentences in Norwegian, and answering some questions about your language learning (can be answered in Norwegian, Danish or English). Each recording will last 20-30 minutes. The recording sessions will occur at your school and at a time that is convenient for you.

The audio recordings will be transcribed, and later be presented to Norwegian speakers as part of the analysis process. They will not know your name, age, or any identifiable information, other than hearing your voice, and be told you are not a native speaker of Norwegian.

The data and information will be collected, and they will not be used for any other purpose than this project. All data and information will be anonymised, and it will not be possible to track the information to the participants in the published results.

Participation is voluntary

Participation in the project is voluntary. If you chose to participate, you can withdraw your consent at any time without giving a reason. All information about you will then be made anonymous. There will be no negative consequences for you if you chose not to participate or later decide to withdraw.

It will not affect your relationship to your school or teacher, or any grades in the classes being observed.

Your personal privacy – how we will store and use your personal data

We will only use your personal data for the purpose(s) specified here and we will process your personal data in accordance with data protection legislation (the GDPR).

- Only the supervisor and the student conducting this project will have access to the information
- Name and contact information will be encrypted and kept separate from other general data. The data will be encrypted and kept on a separate memory card/stick/hard disc which will be kept locked away when not in use
- All names in the project will be replaced by coded names consisting of a letter and a number (you will know which code name is yours. This information will not be shared with others by the MA student or supervisor)

- Information that will be used in the MA thesis will be your gender, your language background (your first language), and ongoing education, since the latter is central to the MA thesis (name of education institution will not be used)
- After the end of the project, all personal information will be anonymised, where all names will be permanently replaced by coded names. The anonymized data will be stored for potential future research. Intended end date for the project is 15.05.2023

What gives us the right to process your personal data?

We will process your personal data based on your consent

Based on an agreement with The University of Oslo, Data Protection Services has assessed that the processing of personal data in this project meets requirements in data protection legislation.

Your rights

- So long as you can be identified in the collected data, you have the right to:
    - access the personal data that is being processed about you
    - request that your personal data is deleted
    - request that incorrect personal data about you is corrected/rectified
    - receive a copy of your personal data (data portability), and
    - send a complaint to the Norwegian Data Protection Authority regarding the processing of your personal data

If you have questions about the project, or want to exercise your rights, contact:

- MA student of this project, Åshild Løvvig via University of Oslo: [redacted]
- Supervisor Peng Li and Haley De Korne, via University of Oslo: [redacted], [redacted]
- NSD (Norsk Senter for Forskningsdata AS), email personvernombudet@nsd.no or call +47 55 58 21 17
- Our Data Protection Officer: personvernombud@uio.no

If you have questions about how data protection has been assessed in this project, contact:

- Data Protection Services, by email: (personverntjenester@sikt.no) or by telephone: +47 53 21 15 00.

Yours sincerely,

Project Leader , MA student *Åshild Løvvig*

Consent form

I have received and understood information about the project Pronunciation in second language and second dialect: The case of Danish learning Norwegian and have been given the opportunity to ask questions. I give consent:

- ☐ To be recorded (audio only) by the MA student, reading a written text. The recording will be anonymised, transcribed and presented to native speakers of Norwegian
- ☐ To that information about my language background and ongoing education will be published anonymously
- ☐ To be observed during language classes classes
- ☐ To participate in a short interview about my experiences of language learning

I give consent for my personal data to be processed until the end of the project.

(Signed by participant, date)

Appendix B

Questionnaire and interview
(Originally in Norwegian)

QUESTIONS ABOUT LANGUAGE CLASSES

PART 1

This is a short interview about language classes and your perceptions and opinions about these. The first questions are open answer, and there are no wrong answers!

- What do you think is the hardest thing about learning Norwegian? (the sounds, the flow, wanting to sound Norwegian enough etc.)
- What do you feel help the most in learning Norwegian, in general? (practice with other Norwegians, talk to yourself, do tongue twisters etc.)
- What kind of activities do you do in language class? Describe them
- What do you feel help the most in learning Norwegian in language class?

PART 2

This is a small survey about your language tuition. There are no wrong answers! If you are uncertain, just put your answer as don't know/not sure

In this survey, the term «action» *(NOR: «handling»)*

WHAT IS AN ACTION?: It can be anything from a physical action, something you do, a gesture, to lean forward or backward to putting emotions, subtext or meaning into the sentence you're saying.

How often do you perform actions or read with subtext when you practice pronunciation in class?

- Very often
- Often
- Don't know

- Not often
- Never

Do you feel comfortable doing actions when you read/pronounce practice sentences in class?

- Very comfortable
- Somewhat comfortable
- No specific feeling about it
- A little uncomfortable
- Very uncomfortable

To what degree do you feel it helps your own pronunciation to interact with your teacher/your fellow students?

- It helps a lot
- It helps sometimes
- Not sure
- It does not help a lot
- It does not help at all

How often do you use/speak Norwegian outside of scool/class?

- Very often
- Often
- Don’t know
- Not often
- Never

To what degree do you feel it helps your own pronunciation to perform actions yourself?

- It helps a lot
- It helps sometimes
- Not sure
- It does not help a lot
- It does not help at all

Appendix C

Interview

Norwegian	English
Hva synes du er vanskeligst med å lære norsk?	What do you think is the most hard part of learning Norwegian?
Hva synes du hjelper mest for å lære norsk generelt?	What do you feel, generally, helps you the most in learning Norwegian?
Hva slags aktiviteter bruker dere i språktimene? Beskriv dem	What kind of activities do you do during language class? Describe them
Hva synes du hjelper mest for å lære norsk i språktimene?	What do you feel helps you the most in learning Norwegian in language class?

Appendix D

Sentences used in recording sessions

Norwegian Sentences	English Translation	Target phonemes in IPA, target sounds highlighted
Les opp disse setningene, med konteksten/underteksten som står under hver setning	Read out loud these sentences with the context/underlying meaning that is written beneath each sentence	-
Jeg klarer ikke dette lenger *Du er lei situasjonen*	I can't do this anymore *You're sick of the situation*	/ʈ/, /ɾ/ k**l**arer, lenge**r**
Jeg så at du kysset henne *Du er sint på kjæresten din*	I saw that you kissed her *You're angry at your boyfriend/girlfriend*	/ʉ/, /y/ d**u**, k**y**sset
Jeg vet at han eier et kart over hele byen *Du har en lur idé*	I know that he owns a map of the whole city *You got a clever idea*	/y/, /ʈ/,/ɾ/ eie**r**, kar**t**, ove**r**, b**y**en
Han burde gå først, han kjenner skogen godt *Du er redd for skogen*	He should go first, he knows the forest well *You're scared of the forest*	/ʈ/, /ɖ/, /ɾ/ bur**d**e, førs**t**, kjenne**r**
Jeg vil at du skal gå, fordi du skremmer meg *Noen gjør deg nervøs/redd/utrygg*	I want you to leave, because you're scaring me *Somebody makes you nervous/scared/feel unsafe*	/ʉ/, /ɖ/, /ɾ/ d**u**, for**d**i, skremme**r**
Sola skinner, som øynene dine *Du vil si noe vakkert til kjæresten din*	The sun shines, like your eyes *You want to say something beautiful to your boyfriend/girlfriend*	/y/, /ʈ/, /ɾ/ so**l**a, skinne**r**, ø**y**nene
Du høres ut som et dyr *Du synes dette er morsomt*	You sound like an animal *You think this is funny*	/y/, /ʉ/, /ɾ/ d**u**, **u**t, d**yr**
Du tar alltid hennes parti *Du er frustrert på bestevennen din*	You always take her side *You're frustrated at your best friend*	/ʉ/, /ɾ/, /ʈ/ d**u**, ta**r**, par**t**i
Han ville så gjerne takle situasjonen, men klarte det ikke *Du er oppgitt over noe*	He really wanted to handle the situation, but he couldn't do it *You're upset about something*	/ʉ/, /ʈ/, /ʈ/ tak**l**e, sit**u**asjonen, k**l**ar**t**e

Hun hadde de styggeste gardinene jeg hadde sett *Du gjør narr av noen*	She had the ugliest curtains I had ever seen *You're making fun of somebody*	/y/, /ʉ/, /ɖ/ h**u**n, st**y**ggeste, gar**d**inene
Fordelene er mye sterkere enn svakhetene *Du er ivrig og glad*	The advantages are much greater than the weaknesses *You're eager and happy*	/y/, /ɖ/, /ɾ/ for**d**elene, e**r**, m**y**e
Han ser ikke sin egen verdi *Du synes dette er trist og dumt*	He doesn't see his own worth *You think this is sad and regretful*	/ɖ/, /ɾ/ se**r**, ver**d**i
Vi har ikke lenge igjen, klokka tikker *Du har fryktelig dårlig tid*	We don't have much time left, the clock is ticking *You're really running out of time*	/ʈ/, /ɾ/ har, k**l**okka, tikke**r**
Jeg hater det ekle gliset ditt *Du sier endelig det du virkelig mener*	I hate that nasty grin of yours *You finally express your true opinion*	/ʈ/, /ɾ/ hate**r**, ek**l**e, g**l**iset
Jeg så en fyr i køen, og han ligner sykt på broren din *Dette er veldig mistenksomt*	I saw a guy in the line, and he looks so much like your brother *This is very suspicious*	/y/, /ɾ/ f**y**r, s**y**kt
Jeg hører kun på folk som behandler meg med respekt *Du er avvisende*	I only listen to people who treats me with respect *You're dismissive*	/ʉ/, /ʈ/, /ɾ/ høre**r**, k**u**n, fo**l**k, behandle**r**