Perceptual Acquisition of Norwegian Close Rounded Vowels by Mandarin Chinese Learners of Norwegian

Cecilie Slinning Knudsen

Master's thesis in Linguistics

Department of Language and Communication Studies

Faculty of Humanities

Norwegian University of Science and Technology

July 2013

Supervised by: Dr. Jardar Eggesbø Abrahamsen

Abstract

This thesis examines the perceptual acquisition of the Norwegian close rounded vowel inventory by six Mandarin Chinese learners of Norwegian. While Mandarin Chinese only has the close rounded contrast /y/ - /u/, Norwegian has the close rounded contrasts /y/ - /ʉ/ and /ʉ/- /u/. Through perception tests of native Norwegian and native Mandarin Chinese informants, problematic areas in this acoustic space are uncovered. The six Mandarin Chinese learners show that acquisition of both contrasts is problematic. The results are discussed within the framework of Optimality Theory, following the basic mechanisms of the Gradual Learning Algorithm (Boersma et al. 2003) and the concept behind Escudero's (2005) L2LP model. What is found is that the Mandarin Chinese learners show tendencies of what may be argued to be conscious knowledge overriding phonological knowledge in the acquisition process. Knowledge about the Mandarin Chinese's tendency to overcompensate when faced with the new category /ʉ/ in Norwegian can aid learners and teachers in the acquisition process.

Acknowledgments

For the things we have to learn before we can do them, we learn by doing them. -Aristotle

While writing this thesis, I ventured into many unknown territories and I have many people to thank for helping me wade through them. Firstly, a huge thank you to Professor Wim van Dommelen for teaching me how to use Praat to manipulate vowels and offering invaluable insights into the phonetic part of this thesis. A thank you also goes to Egil Albertsen who participated in an early stage of brainstorming, and to Christian Dillner Hagen for help on the technical part of my recordings.

I want to thank my father for offering me advice, love and support along the way. I thank my mother for the gift of loving language and for endless hours of discussions about both this thesis and earlier studies. I know our conversations this spring have consisted of mostly crazy ramblings and a barrage of Hertz values and constraint rankings, so thank you for pretending that I made sense. I would also like to thank my brother, Eirik, for doing his fair share of proof reading. Telling me that he found my thesis interesting when I thought no one but me and my supervisor shared that opinion, helped boost my motivation and confidence towards the end.

Signe Rix Berthelin, I have no idea how you managed to put up with me this last year. You kept me sane and helped me clear my head innumerable times. A thank you also goes to Pia Farstad Eriksen, who listened to my many conundrums and offered good advice. To my best friend, Terje Hessen, thank you for being my anchor to the world outside this thesis. My boyfriend, Rene Haugsdal, deserves a thank you for making sure I ate something and took some breaks in front of the TV with him. Thanks to all of you for keeping me calm through the many thesis crises, and letting me ramble on and on for ages about problems I encountered along the way.

My informants! Without you, this thesis would not have been very interesting. Thank you for your interest, and not to mention clicking through hundreds of sound clips. I hope it was interesting to you, too! I cannot express enough gratitude towards Wang Qingqing, who went out of her way to help me gather informants.

Mormor and Moffen, this thesis is in no small part dedicated to you. Your positive influence and support through my academic career, and life as a whole, has been instrumental. I'm grateful to have you in my life.

Last, but by no means least, a great thank you to my supervisor, Jardar Eggesbø Abrahamsen. Our discussions and your comments have been helpful beyond words. Not to mention your replies to my e-mails at every odd hours of the day – science never sleeps! When I started my Master's Degree back in August 2011, I had barely heard about phonology. You ignited my interest and maintained it through two years – and hopefully will continue to do so for many years to come. Thank you for believing in me.

Contents

Chapter 1 Introduction

In today's globalized world, the acquisition of second languages is an increasingly interesting field of study. We not only travel more to distant countries, we also move there, and the need to be able to communicate across national borders is an important issue for many people today. Norwegian might not be a world language, but students come from all over the world to study in Norway. The third[1] largest group of exchange students to NTNU[2] is the Chinese, and several of these students choose to study Norwegian when they arrive in Norway. There are many differences between Norwegian and Mandarin Chinese, and one interesting aspect is how the close acoustic space differs in the two languages. To begin with, where Mandarin has three vowel categories in this space (/i/, /y/, /u/), Norwegian has four (/i/, /y/, /ʉ/, /u/). Additionally, the categories have different phonetic realizations, despite them having the same phonemic labels according to UPSID, (see section 1.1 about IPA and UPSID). This creates a learning task for the Mandarin Chinese that is both phonetic and phonological in that new categories have to be both created and adjusted when learning Norwegian.

One framework for analysis is found in Optimality Theory (OT) where the Gradual Learning Algorithm (Boersma & Hayes, 2001) offers a well-documented method of analysis of acquisition processes (see chapter 4). Optimality Theory (Prince & Smolensky, 1997, 2008) is a linguistic theory that is based on the concept of constraints rather than rules. One of its central principles is that all constraints are active in every language, and that it is the ranking of these constraints that constitute the differences between languages. This means that in acquisition of a new language the learner has to re-rank the constraints of his mother tongue to acquire the ranking of the foreign language. The Gradual Learning Algorithm (GLA) is an OT algorithm which both promotes and demotes constraints in small steps, according to the perceptual learning. The algorithm is claimed to be most accurate when used with Stochastic

---

[1] German (1st) and Spanish (2nd) were not included in this study. An attempt was made to include German, but there were not enough informants to conduct a study. Spanish only has two categories (/i/,/u/) in the close acoustic space, and was therefore left out because the Spanish results could not be directly compared with the results of Chinese Mandarin (or German), whose close space has 3 categories (/i/, /y/,/u/).

[2] Database for statistikk om høgre utdanning, for NTNU in 2011, http://dbh.nsd.uib.no/dbhvev/student/utenlandske_rapport.cfm?vkode=x&brukersort=to&viskode=0&nullvalue=-&landkode=x&studkode=x&progkode=x&semester=1&sti=landkode,studkode,progkode&insttype=x&arstall=2012&instkode=1150&finans=total&fakkode=x&ufakkode=x&beregning=Totalt.antall&valgt_sti=Norges%20teknisk-naturvitenskapelige%20universitet&grupperingstring=a.arstall&sti_hele=instkode,landkode,studkode,progkode&sti_valgt=instkode,landkode,studkode,progkode ,retrieved 15/10/12.

Optimality Theory, a variant of OT where constraints are not fixed, but ranked on a continuous scale (Boersma, 1997). This also allows for optionality in choosing the optimal candidate. The learning models will be presented and discussed in chapter 4.

This study concerns perception only, and aims to establish perceptual category boundaries for the L1 (Mandarin Chinese) categories and the L2 (Norwegian) categories for the informants of the present study. Identification of both L1 and L2 boundaries enables identification of possible problems of acquiring new categories. Furthermore, by examining both beginners and more advanced students of 'Norwegian as a second language', we should be able to see how and where these boundaries shift throughout the learning process. The data in the present thesis, presented in Chapter 3, indicate that there is a change in the boundary between /y/ and /ʉ/ through learning, but not so much when it comes to the boundary between /u/ and /ʉ/ (see section 5.3).

As with children, speech perception is usually ahead of production when learning an L2, meaning that even though a language learner is unable to pronounce contrastive segments, she can be able to perceptually distinguish them from one another (Ashby & Maidment 2005: 184). In perception it is crucial to be able to distinguish between segments that are used to signal semantic differences. Looking at the acoustic data presented in section 1.1.2, a likely obstacle for Mandarin Chinese speakers is to perceive the Norwegian segments /y/, /u/ and /ʉ/ correctly and thereby be able to contrast between minimal pairs such as /ly:s/ ('light'), /lʉ:s/ ('lice') and /lu:s/ ('(marine) pilot'). Theories concerning perception are discussed in chapter 4.

Seeing as the studies presented in section 1.1 use different means to reach their conclusions, and concern production values only, an independent perceptual analysis is necessary to extract the perceptual boundaries between categories. Therefore, those of my informants who are native speakers of Mandarin Chinese will be tested in perception of both L1 and L2 category boundaries. This enables an OT analysis of initial L1 ranking and following L2 reranking of constraints at their current stage of learning. The aim of this study is not to establish identification of 'perfect' categories, nor give an overview of Mandarin Chinese L1 or L2 phonology, but to extract values that can be generalized to be used in a phonological analysis.

The focus will in the present thesis lie on the second formant, F2. In chapter 2, details about why F2 is in focus are discussed. Moreover, how the tests were made, problems that arose while preparing the tests and the execution of the experiments are topics of chapter 2. The methods of creating the stimuli and the process of carrying out the experiments have been allotted sizable space in this thesis, and the present thesis is as such also a methodological study.

Chapter 3 presents the results of the experiment, and these results are discussed within the framework of Stochastic Optimality Theory and the Gradual Learning Algorithm in Chapter 4. Chapter 5 looks at recurring perception patterns shown in the informant replies, and analyzed according to the theories presented in Chapter 4. Lastly, Chapter 6 discusses the results and analyses in light of hypotheses proposed at the end of the present chapter. This chapter will also include a discussion on how the results of this study can be a resource for second language teaching, in addition to possible future research questions.

1.1 A presentation of the vowel systems

When presenting vowels to language students, The International Phonetic Alphabet (IPA) is a well-known and well-used tool, especially in textbooks. It is a useful source for most phonetic purposes, and it can be a powerful resource for students and teachers alike. What this alphabet is not *fully* sufficient for, however, is second language acquisition. This is because the IPA in itself, as an alphabet, does not provide an overview of language specific sounds, but rather offers a framework of reference points to which the symbols can be used to *represent* a sound. The following is to some extent well-known material taken from *The Handbook of the International Phonetic Association* (1999). It is included here to serve as a background for the subsequent discussion.

The International Phonetic Association favors generalizations in their Phonetic Alphabet so as to allow "for a very economical analysis of the complex and continuously varying events of speech", (p. 6) in a manner "that it is widely understood", (p. 30). The segments, described after production and auditory characteristics, are to be seen as 'target' descriptions or reference points. For vowels, which are the topic of this study, the notion of 'target' descriptions or reference points is of special importance as the vowel space is continuous.

Vowels are represented in the 'Vowel Quadrilateral', described as "an abstract vowel space" (p. 10), and was first created by Daniel Jones as a visual aid to see how the vowels are articulated. However, phoneticians today see it as rather representing the auditory space. The Quadrilateral has the parameters 'close' – 'open' and 'front' – 'back'. When a vowel is 'close', the tongue is near the roof of the mouth. A vowel is described as 'open' when there is space between the tongue and the roof of the mouth. A 'front' vowel is pronounced when the highest point of the tongue is at the front of the area where vowel articulations are possible. Conversely, a ‘back’ vowel is produced when the tongue is at the back of the mouth.

Based on these criteria, we get cardinal vowels: The extreme vowels that are maximally open, close, front and back: [i, u, a, ɑ]. The remaining vowels are defined based on auditory spacing, where the differences between each vowel and the next in the series are auditorily equal. Consequently, the Quadrilateral is not exclusively based on articulation and therefore not an accurate representation of vowels in use. The cardinal vowels have 8 primary cardinal vowels and 8 secondary cardinal vowels. The latter are the rounded counterpart to the primary vowels and always shown to the right in the Quadrilateral. In addition, there are two secondary cardinal vowels in the mid-open area, as well as vowels for the mid-central area and intermediate positions.

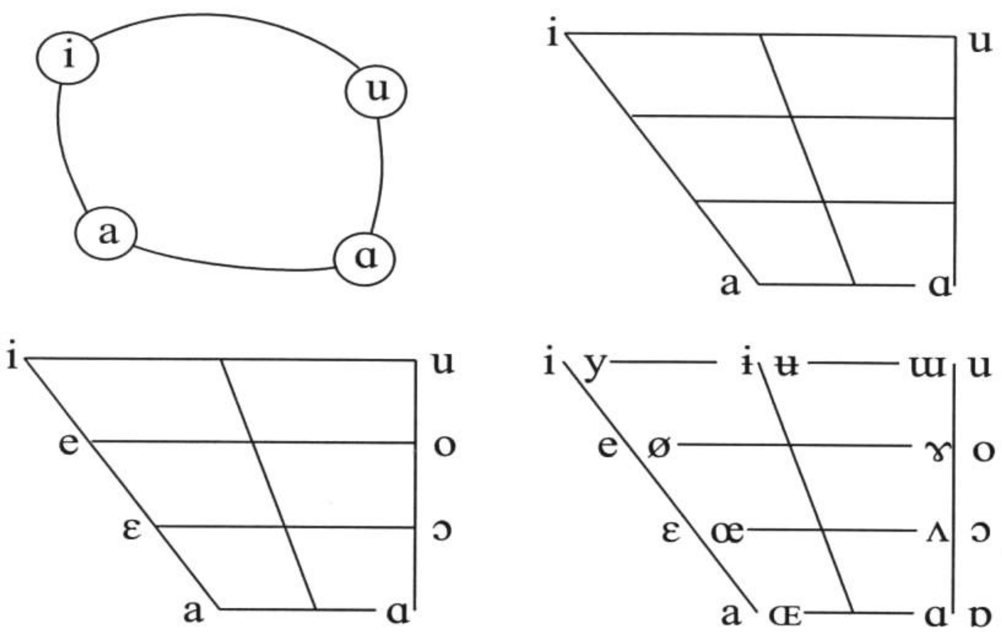

Figure 4 The vowel quadrilateral and cardinal vowels. Above, the relation between the vowel quadrilateral and the vowels shown in figure 3; below, the primary cardinal vowels and all cardinal vowels.

Fig 1.1 The Primary and Secondary Cardinal Vowels, including original figure caption, from The Handbook of the IPA (1999: 12).

This ‘continuous vowel space’ leaves more room for variation in pronunciation of vowels than for consonants; a labiodental segment has a narrower place of articulation than a close-back vowel. As a result, the description ‘close-back’ in the phonetic and phonological literature encompasses a range of vowel qualities in the proximity of the cardinal vowel [u]. Most languages "use vowels which are similar to, but not as peripheral as, the reference points", (p. 13). To achieve a detailed phonetic description using IPA, one can use diacritics. This, however, might be argued to defeat the purpose of the IPA as mentioned above: An economic analysis that is widely understood.

What if we rather complemented the IPA with language specific Quadrilaterals? Much of this work has already been done by phoneticians, where they measure the Hertz values of the formants of the vowels and plot the segments on their exact acoustic space in the Quadrilateral. We have three examples of work that measure the formants of vowels in the present thesis, van Dommelen (p.c.), Zee & Lee (2001) and Pätzold & Simpson (1997).[3] Nevertheless, in large language projects like UPSID, the reference points of the IPA Quadrilateral are the only ones referred to.

UPSID, the UCLA Phonological Segment Inventory Database, was collected by researchers at the University of California under the supervision of Ian Maddieson in the 1980s. This database contains data on the phonological systems of 451 languages.[4] UPSID is based on the a priori cardinal vowels of the IPA, but is in itself of an a posteriori nature as it is a depiction of empirical data, with a typological aim. Maddieson (1984), as IPA, aims to "provide a reliable basis for […] generalization", (p. 1), and the observations and hypotheses about phonological universals that follow from this are "relative rather than absolute" (p. 2).

Furthermore, the UPSID database is a collection of segment inventories from different individual sources. This can mean discrepancy between the degree of phonetic detail each inventory contains because it "depends greatly on the phonetic judgments and transcription methods of the field linguist" (1984: 138). The parameters used for vowel description in UPSID constricts to height, backness and lip-rounding (1984: 123). Maddieson (1984) notes that vowels, in this case mid vowels, might not be described accurately or with elaborate specification (p.123). This in itself is not surprising, as the individual field linguist was mapping within a single language and did not necessarily see the need for more information than what was needed to distinguish between the phonemes of that particular language.

[3] See also Kristoffersen (2000) who measured Norwegian formant values (Urban East Norwegian) and plotted the values into a Quadrilateral.
[4] UPSID: http://www.linguistics.ucla.edu/faciliti/sales/software.htm#upsid, retrieved 09.01.13.

UPSID is the foundation for many studies within linguistics, one of them being the Computer-Assisted Listening and Speaking Tutor (CALST) used in the courses in Norwegian for Foreigners at NTNU.[5] Their L1-L2map provides “a tool for contrastive analysis of the phonetic segment inventories” (CALST). The researchers behind CALST have extended the UPSID data by adding positional information for consonants, but no additional work has so far been done on vowels. The positioning of language-specific vowels in the L1-L2 map quadrilateral is therefore exclusively based on UPSID data and plotted on the cardinal vowels of the IPA.

In my opinion, this can lead to misunderstandings and confusion for learners of a second language. One symbol in the Quadrilateral can represent two language-specific sounds that are different to such a degree that either mispronunciation or misunderstanding can occur. IPA’s aim is not to be language-specific, but problems arise when IPA symbols are used for pedagogic purposes. [y] is a good example here, as the pronunciation of a German /y/ is not the same as that of the Norwegian /y/. These two vowels are nevertheless considered equal in CALST (shown by green coloring in figure 1.2 below) and represented by the same symbol, /y/, in IPA. If the German learner then follows the IPA, or CALST, in acquiring the Norwegian version of this segment, she *will* get it wrong. Consider Figure 1.2 below, retrieved from the CALST L1-L2 mapping tool in October 2012:

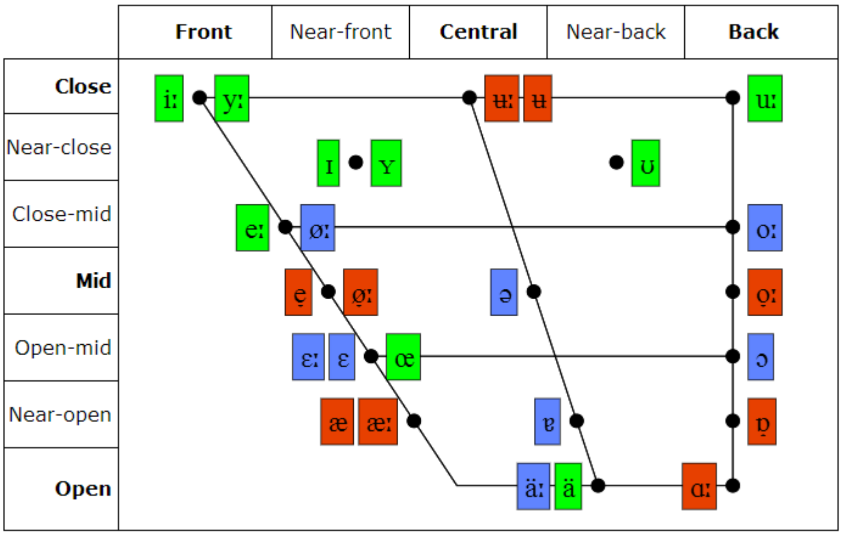

Fig. 1.2 A comparison of German and Norwegian vowels. German is represented in blue, Norwegian in red, and green is for overlap between the two languages. (CALST[6])

[5] CALST: http://www.ntnu.edu/isk/calst-for-learners
[6] L1-L2 Map, http://calst.hf.ntnu.no/L1-L2map, CALST, NTNU, Retrieved 10.10.12.

The International Phonetic Association (1999) does state that the IPA does not provide language-specific phonological analyses (p. 30), but it is also specified that helping “learners of foreign languages with phonetic transcriptions to assist them in acquiring the pronunciation” is one of IPA’s primary goals, (1999: appendix 1). Moreover, CALST is designed for the sole purpose of teaching Norwegian as a second language. The differences between segments may well be addressed otherwise in the course, but the CALST Quadrilateral is misleading, presenting the Norwegian and the German [y] as identical segments. Consider figure 1.3 below[7] where I have plotted the Norwegian and German close vowels according to the formant values from Norwegian formant values from van Dommelen (personal communication, Norwegian – see section 1.1.1 below) and Pätzold and Simpson (1997, German):

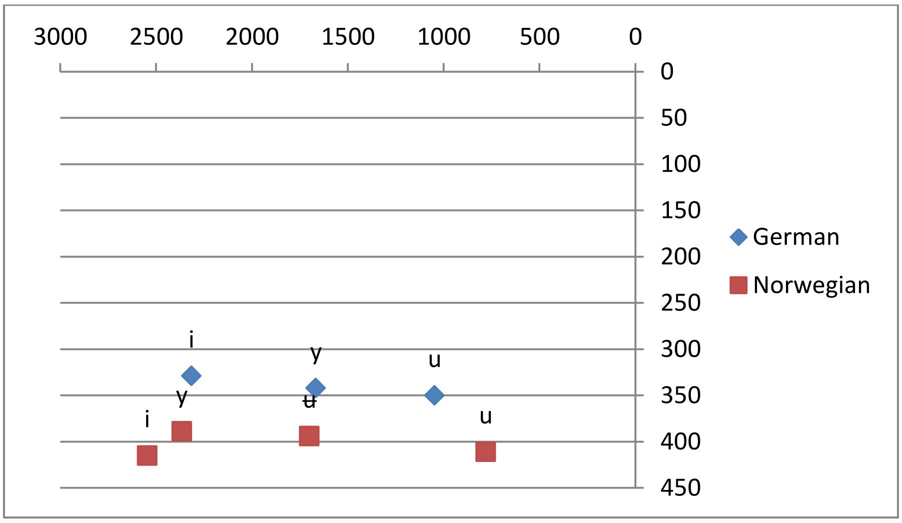

Fig. 1.3 A comparison between Norwegian and German close vowels in terms of F1 and F2 based on van Dommelen (p.c., Norwegian) and Pätzold and Simpson (1997, German).

Here we can see a clear discrepancy between the languages, recognizing immediately that one cannot rely on IPA, and in turn UPSID, categorization while teaching or learning languages.

To achieve a sufficient and uniform description, adaption of segments rather than phonemes for databases seems preferable. I suggest adding values such as [F1: 350, F2: 2600, F3: 4700] to the description of vowels for such databases, as we are now in possession of technology that enables easy processing and analysis of segments based on Hertz values. Such features

[7] Made in Excel after the instructions of http://www.indiana.edu/~l541/week%205/Creating%20a%20vowel%20system%20in%20Excel.pdf, retrieved 15.03.13, using xy chart labeler

enable phonological analyses of category boundaries, where constraints of the type '[2600 Hz] not /u/' (see section 4.2) in Optimality Theory are employed.

The difference between the Norwegian and Mandarin Chinese close vowels are not as severe as the difference discussed above. However, by taking a closer look at the perception of formant values we can get valuable insight into what areas the Mandarin Chinese students of Norwegian are more likely to experience problems acquiring.

1.1.1. Norwegian

Norwegian[8] has one of the more complicated language systems with 16 vowels in a rectangular shape, 18 if we include the long/short variations of /ʉ/ and /æ/. According to Husby & Kløve (2001), the most common language systems are triangular with 5 to 7 vowels (42ff). From an IPA and UPSID point of view, the Norwegian quadrilateral as represented by CALST looks as in figure 1.4 below:

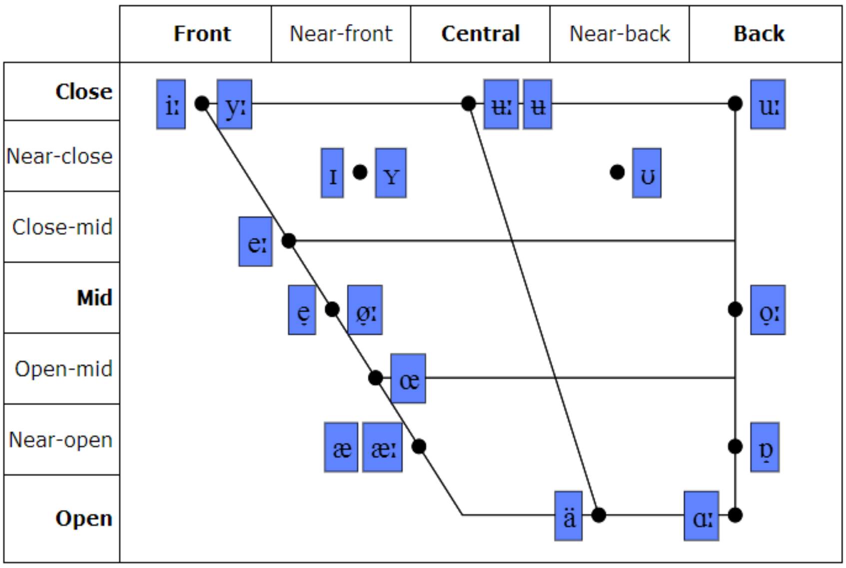

Fig 1.4 The Norwegian vowel inventory (CALST[9]).

The short allophone of the Norwegian /i/ is here represented by the lax vowel /ɪ/ and the Norwegian short allophone of /y/ as the lax vowel /ʏ/. Norwegians can have less or more vowels than this due to dialectal differences, so my informants for the Norwegian category boundaries were chosen from different parts of the country to see if dialect and region had any effect on perception (see chapter 3 for results).

[8] It is worth noting that the Norwegian dialects differ in what categories they have. The dialect is not specified in this case, but is assumed to be standard Urban East Norwegian.
[9] L1-L2 Map, http://calst.hf.ntnu.no/L1-L2map, CALST, NTNU, retrieved October 2012.

The close, rounded vowels are the topic of this thesis, and in Norwegian these are /y/, /ʉ/ and /u/. The Norwegian /y/ differs from other language's /y/ (Mandarin Chinese included) by its protrusion. In an unpublished study on Norwegian vowel quality, van Dommelen (p.c.) measured the formant frequencies of 3 Norwegian males and 3 Norwegian females, and calculated the mean values of the male and female responses. By studying these formant values, we get a more detailed picture of the Norwegian vowels. The results from the Norwegian female speakers are presented in Table 1.1:

Segments	F2 value	F3 value
/iː/	2547	3174
/yː/	2367	2957
/ʉː/	1707	2468
/uː/	781	2754

Table 1.1 Formant values of female native speakers of Norwegian (van Dommelen, p.c.).

Judging from F2, this tells us that the /ʉ/[10] is acoustically closer to /y/ than to /u/ in Norwegian. An interesting question here is exactly how large the acoustic space of /ʉ/ is. The results from the perceptions tests, shown in Chapter 3, will shed light on this question.

1.1.2. Mandarin Chinese

Mandarin Chinese is made up of many dialects that can be quite different, and speakers from different parts of the country might not understand each other. This thesis will focus on what is known as Standard Chinese, and the literature used is that on Beijing Mandarin. This dialect is close to identical to Standard Chinese.

According to CALST,[11] the Mandarin Chinese system is what Husby & Kløve (2001:42ff) refer to as a common language system with its triangular shape of 6 vowels: 3 close vowels, 2 mid vowels and 1 open vowel. The Mandarin Chinese close perceptual space thus only consists of one third of the amount of what the Norwegian does, and there are only two close rounded categories in Mandarin Chinese: /y/ and /u/.

Four tones can be applied to any Standard Chinese vowel, and these tones carry contrastive meaning when applied to vowels (ibid: 36). In this study, the flat tone, 1, is used.[12] Figure 1.5

[10] All the Norwegian vowels discussed in this thesis are long, and will from here on out not be transcribed with the marker of length, [ː].
[11] L1-L2 Map, http://calst.hf.ntnu.no/L1-L2map, CALST, NTNU, retrieved 10.10.2012.
[12] As with Norwegian length, the Mandarin Chinese vowels will not be transcribed with the marker for tone 1.

below is a CALST L1-L2 mapping between Norwegian and Mandarin Chinese vowels. Red symbolizes Norwegian, blue symbolizes Mandarin Chinese, and green is overlap between the languages. This figure is also based on UPSID data and gives a generalized picture of the Mandarin Chinese vowels:

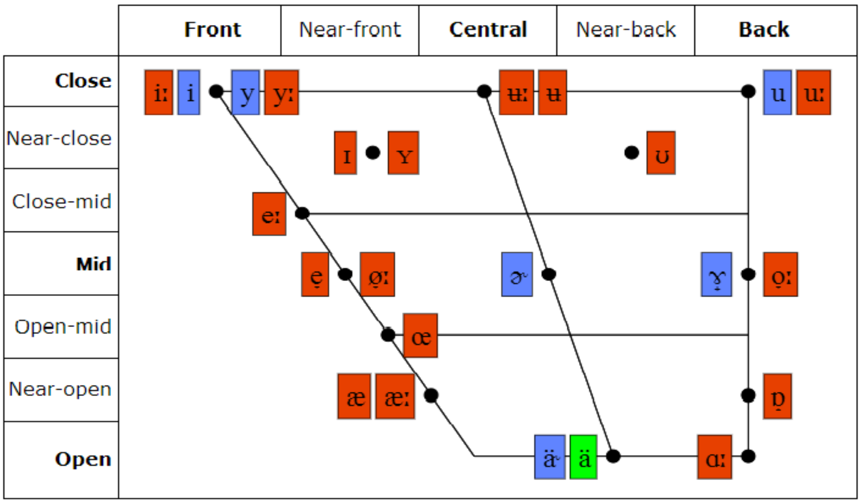

Fig 1.5 A comparison of Mandarin Chinese and Norwegian vowels. Mandarin is represented in blue, Norwegian in red and green is for overlap between the two languages. (CALST[13])

In the figure above, we see that the feature that separates the Norwegian and Mandarin close back and close front categories is that of length. The figure implies that this is the only difference between the categories, and as we shall see this is also correct in terms of F2 production.

In their article, ”An Acoustical Analysis of the Vowels in Beijing Mandarin”, Zee and Lee (2001) present a spectral analysis of the 6 vowels in Beijing Mandarin (BM). Through analysis of the recordings of 20 native speakers of BM, 10 males and 10 females, they extracted the mean second formant values of each vowel. The table below renders the results from the close vowels of the female speakers:

[13] L1-L2 Map, http://calst.hf.ntnu.no/L1-L2map, CALST, NTNU, retrieved 10.10.2012.

Female	F2		F3	
	mean	s.d.	mean	s.d.
/i/	3036.76	185.03	3847.56	262.88
/y/	2327.36	141.18	2999.88	180.40
/u/	758.68	111.73	3308.82	275. 95

Table 1.2 Average F2 and F3 values (in Hertz; n = 50) and their standard deviations (s.d.) for the vowels [i, y, u], (Zee and Lee 2001: 644).

With this information, we can take a more in-depth look at the contrasts between the languages' close vowels:

Norwegian	F2	Beijing Mandarin	F2
/i:/	2547	/i/	3036.76
/y:/	2367	/y/	2327.36
/ʉ:/	1701		
/u:/	781	/u/	758.68

Table 1.3 A comparison between the close segments of Norwegian and Beijing Mandarin females.

Beijing Mandarin shows a greater difference in F2 values between /i/ and /y/ than Norwegian does for /i/ and /y/. The Mandarin Chinese learners are thus expected to have difficulties with distinguishing between L2 /i/ and /y/[14]. The cluster of L1 and L2 categories in the front close vowel space may evoke problems in the perception of both L2 front, close vowel categories.

More importantly, the differences between the BM /y/ and the Norwegian /y/, in addition to those between the Norwegian /u/ and BM /u/, are seemingly insignificant. This would suggest an easy acquisition of these two L2 categories. The Norwegian /ʉ/ is closer to the Mandarin Chinese /y/ than to Mandarin Chinese /u/, but without knowing the full extent of the category boundaries, this is not sufficient information for making hypotheses on whether a Mandarin native will be more likely to perceive a Norwegian /ʉ/ as an /y/ or an /u/.

[14] The perception of /i/ is not in the scope of this thesis.

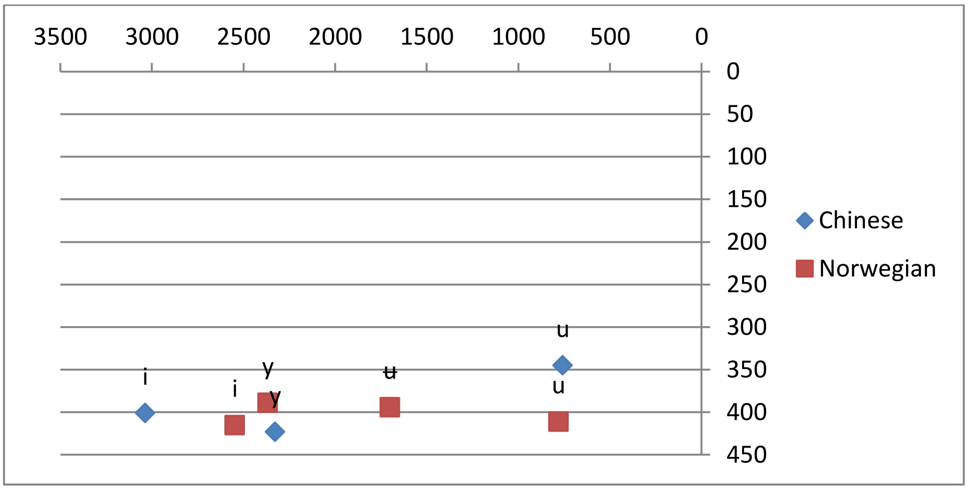

Fig 1.6 A comparison of Mandarin Chinese and Norwegian vowels based on Zee and Lee (2001) and van Dommelen (p.c.). Mandarin is represented in blue and Norwegian in red.

From the acoustic information given in figure 1.6 above, we can predict that the Mandarin Chinese will experience less difficulty establishing a /u/ category in the L2 than establishing the L2 categories that are clustered in the front area. /ʉ/ also seems to be more likely to be assimilated to, i.e. equaled to, the /y/ in the initial L2 stage considering the acoustic distance. This assimilation can mean that both /ʉ/ and /y/ are considered the same category in the L2, or that parts of the /ʉ/ category are recognized as /y/. The rest of the /ʉ/ category, i.e. the Hertz values that constitute a Norwegian /ʉ/, can be established as its own category or assimilated to the /u/. It is important to note that these values are based on production only, and as we will see in later chapters, the perceptual categories have more variation between the languages.

Importantly, in the acquisition process there is a difference between the phonological dimension and the phonetic dimension. The phonological dimension is the categories in a language's vowel inventory, while the phonetic dimension is the realizations of these categories. The learning task in acquiring an L2 phonology is thus to determine the phonetic boundaries of the phonological categories. From what is seen in figure 1.6 above, both the Norwegian and Mandarin Chinese categories that are labeled /y/ in the IPA seem to belong to the phonological category front.[15] Similarly, the /u/ categories of both languages seem to belong to the phonological category back.

If that is the case, the Mandarin Chinese have the task of acquiring only one new phonological category, central, or as named in the IPA, /ʉ/. For /y/ and /u/, the learning task for the

[15] A complicating factor to this assumption may be that Norwegian/y/ and /u/ are specified as long, while the Mandarin Chinese /y/ and /u/ are not. This issue is not investigated in the present thesis.

Mandarin Chinese is thus of a phonetic nature, where the category boundaries are to be adjusted from the L1 category to the L2 category. The new /ʉ/ category also needs to be adjusted, given that it is successfully established in the L2 category inventory and not seen as a bad example of either /y/ or /u/, or both.

In the present thesis, the phonological categories are discussed under the labels given by the IPA, i.e. front is referred to as /y/[16], central as /ʉ/, and back as /u/. The Mandarin Chinese front category will be referred to as "L1 /y/" and the Norwegian front category as "L2 /y/". Similarly, the back category is referred to as "L1 /u/" for Mandarin Chinese, and "L2 /u/" for Norwegian.

Despite the F2 values being so similar, a mere intuitive perceptual approach tells the listener that a Mandarin Chinese /y/ and a Norwegian /y/ are not identical. The reason for this is probably lip protrusion. The Norwegian /y/ is protruded, while the Mandarin Chinese is not. Protrusion will normally lower all formant values, but especially affect F3 (Asbhy & Maidment 2005: 74). This possibility is not investigated further in this thesis as the focus lies on F2 (see also discussion concerning F3 in section 2.1).

Interestingly, some studies propose that the Mandarin Chinese notation for [y] is replaced with [iu], (Hartmann (1994), Hockett (1947), Martin (1957) and Hsueh (1986), cited in Duanmu 2007: 37). Duanmu (2007) does not pursue this issue further, and the dates of the referred studies can indicate that this view is outdated. However, it is worth noting that this diphthong replacement goes from front to back, through the perceptual space of the Norwegian /ʉː/.

1.2 Hypotheses based on theory and literature

Based on the theory presented in Chapter 4, it is predicted that Mandarin Chinese learners of Norwegian in the beginner stage will show signs of their L1 phonology copying to the L2 phonology. From the literature in the present chapter, this copying is not assumed to pose a problem for the segments /y/ and /u/ as their production values are quite similar in Norwegian and Mandarin Chinese. The question in the present thesis is to what degree these categories differ perceptually.

[16] As there are no unrounded counterparts to the rounded vowels in this thesis, the added information of /rounded/ is usually not included in subsequent discussions.

Furthermore, as the Norwegian category /ʉ/ resides between the L1 /y/ and /u/ categories, it is unclear which category /ʉ/ will map to, or if /ʉ/ will map to both L1 /y/ and /u/. From the acoustic analyses presented above, it is more likely that /ʉ/ is mapped to /y/ than to /u/ since /ʉ/ is analyzed as being acoustically closer to L1 /y/ than /u/. This mapping should theoretically show in the interlanguage states of the beginner students: If they confuse /y/ and /ʉ/, it is likely that the Mandarin Chinese mapped L2 /ʉ/ to L1 /y/, and similarly mapped /ʉ/ to /u/ if they confuse /u/ and /ʉ/ at the beginner stage.

More advanced learners are expected to gradually construct a separate phonology for the new language, moving from an interlanguage state to two separate phonologies for the respective languages. It is thus expected that the advanced learners have achieved the establishment of the new category /ʉ/, and that they have attained category boundary adjustment according to the L2 to a higher degree than the learners at the beginner stage.

Chapter 2 Method

The aim of this thesis is to find out where the informants draw the category boundaries in both their L1 and their L2. In order to identify this, a perception test is required. The results from the tests are translated, within the framework of Optimality Theory and the Gradual Learning Algorithm, to constraints for each of the values in the test, e.g. '[1600 Hz] not /y/' (see Chapter 4 for more on these constraints). The Mandarin Chinese perception test results will inform us of the ranking of these constraints in their native language, and the Norwegian perception test results will show how this ranking has been altered.

Choosing the right method is crucial to extract useful and viable results. This chapter presents and discusses the methodological choices and procedures related to the creation of the perception tests. This study is approved by the Norwegian Social Science Data Services (see appendix 3). All informants signed a consent form and everyone was informed of the purpose and goals of the study in oral and in written form (see appendix 1).

2.1 Phonetics

To create perception tests, knowledge about phonetics and phonetic tools are crucial. The following presentation of vowels in a phonetic perspective is a somewhat superficial account since phonetics is not the primary objective in this thesis. The vowel manipulation procedure is approved by van Dommelen, a professor in Phonetics, and thus the chances that the result is sufficient as a tool for the phonological analysis are increased.

The physiological process of uttering a vowel includes the oral cavity, the lips and the vocal folds in the larynx. By changing the size of the oral cavity and the positioning of the lips, we can articulate different vowels. The acoustic energy coming from the vibration of the vocal folds is called the input or excitation, and some of its component frequencies are "picked out and reinforced by the resonant characteristics of the vocal tract", (Ashby & Maidment 2005: 70). The vocal tract acts as a filter, picking up energy from some frequency regions while leaving others out, and thereby creating an output spectrum. The peaks of energy in this spectrum are called formants (ibid: 71). These formants have frequencies, and it is these frequencies that are measurable in Hertz values and which enable us to distinguish vowels from one another perceptually, (ibid: 71).

Vowels are made up of several formants and the traditional view is that we need at least perceptual information about the three first formants, F1, F2 and F3, to determine what vowel we hear. F1, the lowest formant, is determined by the raising of the tongue and signals vowel height. The closer the tongue is to the roof of the mouth, the lower the frequency. Low F1 frequencies correspond to 'close' vowels, and high F1 frequencies correspond to 'open' vowels, (ibid: 73). In this study, all segments are close vowels, and thus we are more concerned with the values of the second formant. F2 values signal vowel location on a continuous scale from 'front' to 'central' to 'back'. These values are created by the position of the tongue in the mouth, further back for back vowels, up against the soft palate for front vowels, etc. The higher the F2 frequency, the more front a vowel is (ibid: 74f). F3 might also be needed to distinguish between vowels, particularly when it comes to lip rounding.[17] This third factor can lower all formant frequencies. The most notable change is found in the second formant, but the third formant can also be considerably lowered, (ibid: 75).

Vowels are not static entities, but vary throughout production. Deciding what formant values a vowel has is therefore a demanding task, and the results are often imprecise. It is also important to note that "[i]t is the pattern of formant frequencies and their relationship to one another that is important rather than absolute values", (Ashby & Maidment 2005: 72). Ladefoged (2003) suggests determining the formant value by measuring "an interval near the middle of the vowel", (p. 104). The Praat manipulation procedure, however, manipulates the average value of the entire vowel (see section 3.3 below). Consequently, the average value has to be my criterion as well.

According to Ashby & Maidment (2005), "it is rare for a phonetic distinction to be signaled by a single acoustic cue", meaning we may need both F2 and F3 to make a distinction. However, one of them might suffice, and this is called cue redundancy (p. 183). The learning algorithms presented in Chapter 4 only take F2 into account, but some of them also rely on length differences. The second formant signals to what degree a vowel is front, central or back, and is therefore essential to this experiment. Conversely, F1, the formant that signals height, can be disregarded as the vowels in this thesis are all close vowels. The uncertainty lies with the third formant, F3. The recordings (see section 2.3 below) showed that the F3 values were 3490, 2822 and 2668 for Norwegian /y/, /ʉ/ and /u/ respectively. We can see that the greatest difference here was between /y/ and /ʉ/, and thus it was necessary to investigate the importance of F3 on these vowels more closely.

[17] The only unrounded vowel here is /i/, and it was not a part of either of the manipulations.

Consequently, an informal experiment was conducted to determine whether or not it was necessary to manipulate both F2 and F3, or if manipulation of F2 alone was sufficient to create a Norwegian /ʉ/ from a Norwegian /y/. It proved impossible to create stimuli with no manipulation of F3 through the Praat manipulation procedure, using my own recordings. The program alters all formant values somewhat, despite being instructed to only alter F2 (see section 2.3 below for details on the manipulation procedure). However, the changes are miniscule. I started out with a Norwegian /y/ with the original resample values of F2:2811 and F3:3443 and successfully created a Norwegian /ʉ/ with the values of F2: 2054 and F3: 3467. These stimuli were played to my supervisor and a professor of phonetics, both of whom judged the latter stimulus to be a Norwegian /ʉ/. This means that it is inconsequential for the Norwegian perception whether a /ʉ/ has an F3 value of 2822 or 3467, as long as the F2 value is around 2054.[18]

I conducted the same test on recordings of Mandarin Chinese and came to the same conclusion. Here, the difference between /y/ and /u/'s F3 values were 335 Hz. Such small differences in values are not expected to be important for correct vowel perception. Additionally, a gradual manipulation of the Mandarin Chinese F3 values would include 19 intervals of minus 18 Hz, something Praat was unable to accomplish successfully. Based on these experiments I concluded that manipulation of F2 sufficed for this study. The question of F3 will also be briefly addressed in Chapter 4.

2.2 Test words and frame sentences

According to a study on vowels in consonantal context by Kewley-Port (1995), one should avoid minimal pairs containing /m/ and /l/ when testing subjects' ability to resolve formant frequency, (p. 3143). She discovered high between-subject variability when testing their perception of /mɪm/ and a large increase of identification with /lɪl/ compared to isolated /ɪ/ (ibid). As for other consonantal contexts, Kewley-Port claims that they have little effect on F1, but that the frequency resolution for F2 appears to be degraded (ibid). She observes that this has to do with particular consonant stimuli, such as formant transitions, separation of the onsets of the formant transitions and "the duration of the steady-state portion of the vowel" (1995: 3144). Her conclusion is that consonant context has an effect on some informants' ability to correctly identify the vowel, but that it mostly has little effect (1995: 3146).

[18] The original recording of the /ʉ/ showed an F2 value of 2180.

Building on Kewley-Port's findings, the minimal pairs chosen for this study are CVCs with "longer steady-state vowels", (1995: 3145). The vowel categories of this study are therefore long: [y:, ʉ:,u:]. The Mandarin Chinese vowels are not specified for length in UPSID (see Chapter 1). For Chinese, however, tone is an important factor, and it was decided to go with tone 1, a flat tone (see section 2.3). A decision was also made to have open syllables, CV, instead of Kewley-Port's recommended CVC syllables, to further enhance the length of the vowels for manipulative purposes, and to reduce the effect of consonantal context.

One Mandarin Chinese informant, who did not participate in the perception test, recorded the speech signals which were manipulated for the perception test. She also participated in finding native words for the test, where the consonantal context of the words had to adhere to the previously mentioned specifications. Due to the tonal nature of Mandarin Chinese, finding exact minimal pairs proved difficult. Given the instructions of choosing open syllable words, using tone 1 and excluding the aforementioned consonants, the informant told me that the words had to be nonsense words. To avoid different consonantal influence on the vowels, the same consonantal context had to be applied to both languages. The choice fell on the consonant [n]. Fricatives were excluded to avoid prominent external cues to the vowel quality, especially rounding, from co-articulation, and stops were excluded to avoid aspiration. Granted, [n] will transfer some of its nasal features onto the vowel, but it was here deemed the least obtrusive change to the vowel. As a result, the stimuli for both languages went from [nu] to [ny], with all intermediate values included in the continuum.

The OT learning models presented in section 4.2 relies on the assumption that phonological learning follows semantic learning and as such it would have been more accurate to test the informants with actual words instead of nonsense words, where the vowels marked contrastive meaning between them. With the aforementioned criteria, this was simply not possible.

Given that the informants were tested in both L1 and L2, frame sentences proved necessary to 'tune' them in to the correct language they were being tested in at any given time. Placing the stimuli in the middle of a sentence also helps prevent creaky voice and phonetic lengthening. Frame sentences of the type 'What I said was …' were chosen:

Norwegian:	'Det	var	…	jeg	sa'
	Pro	V		Pro	V
	It	*was*		*I*	*said*

Mandarin Chinese:[19] ‘Wo shuo de zhe ge … zi ta bing bu cun zai’

我	说	的	这	个	…	字,	它	并
wo	shuo	de	zhe	ge	…	zi	ta	bing
Pro	V	Part	Det	Cl		N	Pro	Prep
I	*say*	*PART*	*this*	*CL*		*word*	*it*	*at*

不	存	在。
bu	cun	zai
Adj	Neg	V
all	*NEG*	*exist*

2.3 Creating the stimuli

For the perception test, a speech continuum with different formant values was needed. The first step was to record native speech to be used as a basis for the stimuli. The Norwegian and Mandarin Chinese recordings were performed in the soundproof studio at the Department of Language and Communication Studies, NTNU, in December 2012. The equipment used was a Shure KSM44 microphone, and the recordings were saved on a hard drive with the sampling frequency of 44,1 kHz, 16-bit quantization.

After failed attempts to manipulate the Mandarin Chinese vowels, it became clear that a new recording was needed. The reason behind this was that the Mandarin Chinese informant’s /y/ varied greatly in F2 values throughout the pronunciation, and had significant creaky voice in the middle of the vowel, ref. figure 2.1 below.

[19] Informant’s own glossing and annotation

_y__with_tone_3

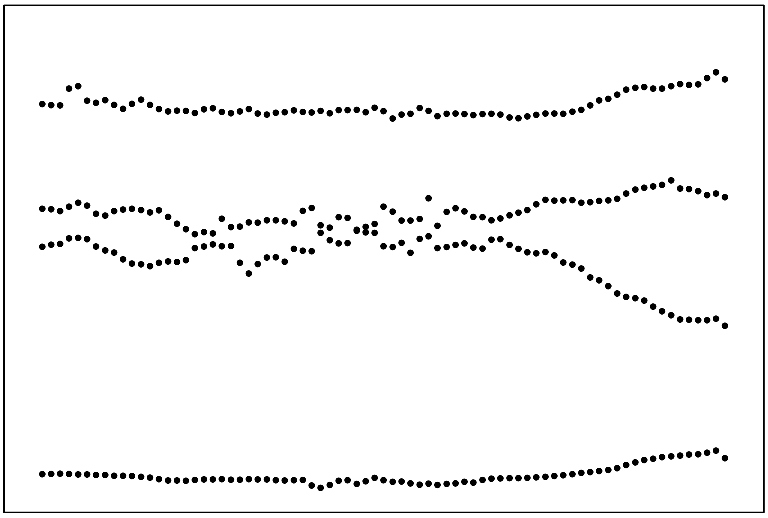

Fig. 2.1 The Mandarin Chinese /y/ pronounced with tone 3.

My Mandarin Chinese speaker and I came to the conclusion that this correlated with the tone used, in this case tone 3, which is a fall-rise tone. The difference can be seen in figure 2.2 below, where a flat tone, tone 1, was used instead:

y_with_tone_1

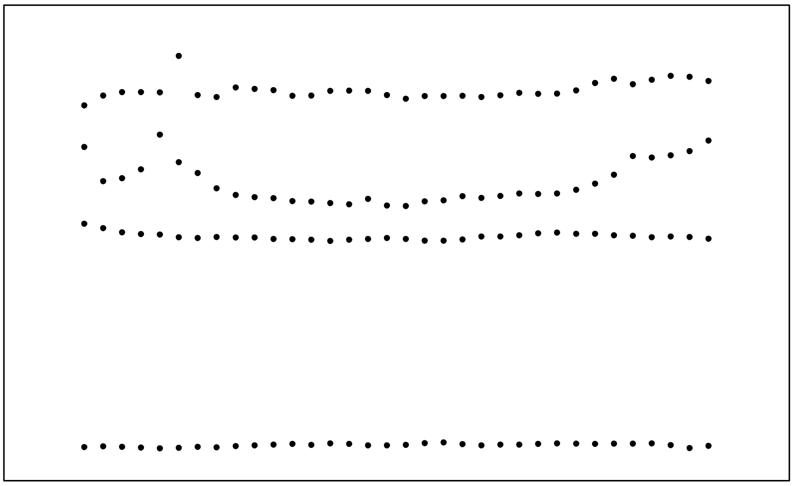

Fig. 2.2 The Mandarin Chinese /y/ pronounced with tone 1.

The new recording was done in January 2012 with the same person as before, but with instructions of using tone 1 and of including more repetitions of the sentences so as to get more vowels to test for manipulation.

As humans are not able to alter acoustic cues with perfect accuracy, the stimuli for the perception tests were created with digital manipulation. The analysis and manipulation of vowels were carried out in the phonetic program Praat (Boersma & Weenink, 2012), and pre- and post-production sound file fitting was done in the audio editing application Audacity (1.3.13-beta, Ash, Chinen, Crook).

A couple of choices had to be made for the Herz values of the stimuli. First of all, was it more important to have an equal amount of stimuli per language, or that the stimuli differed equally in values? The choice fell on the latter because that would give a better foundation for comparison, and thereby the value of 100 Hz was chosen. The end result was 20 stimuli for Norwegian and 19 for Mandarin Chinese, and thus there was not much difference in either values or amount. Unfortunately, as will be discussed in section 2.3.1, it proved difficult to create stimuli with exactly 100 Hz intervals through Praat. Consequently, there is little uniformity between the Norwegian and Mandarin Chinese stimuli. For example, where the Mandarin Chinese continuum had the stimuli 1861 Hz – 1936 Hz – 2074 Hz, the Norwegian continuum had the stimuli 1899 Hz – 1956 Hz – 2053 Hz. Also, the Norwegian continuum ended up with two stimuli around 2000 Hz, and the Mandarin Chinese lacks a stimulus around 2200 Hz. There are, however, intermediate values that make up for these differences.

Next, there was the decision about where to start and end the continuum. There are no vowels residing in a further back environment than the vowels in this test, so it was not possible to create a category boundary on that end of the /u/'s in either language. I would then have to add a "none of the above" option not fitting for a forced alternative test (see section 2.4). As for the values at the start and end point, I used the F2 values from the recordings as reference points. The value for the recorded Norwegian /u/, 781, is identical to that of van Dommelen (p.c., see Chapter 1), while the value for the recorded Mandarin Chinese /u/, around 600 Hz, is about 100 Hz lower than that of Zee and Lee (2001, see Chapter 1). For the Norwegian /y/, van Dommelen measured the average value at 2367, and 2547 for /i/, and the continuum goes beyond these values. Zee and Lee measured the average value for /y/ at 2327, and 3036 for /i/. The Mandarin Chinese continuum was ended where an acceptable /y/ was produced, at 2547, and therefore does not include the values for Mandarin Chinese /i/.

I decided not to test for the boundary between /y/ and /i/ because my hypothesis was that the problem of the Norwegian acquisition by the Mandarin Chinese would lie in the establishment of the /ʉ/ category between /u/ and /y/. In hindsight, this decision was wrong. The results of the tests show that the /ʉ y i/ area is problematic, and the regulations to the test along the way further enhanced the problem of analysis. Adding an /i/ option only for the L2 test of Norwegian vowels for Mandarin speakers and not for the L1 Mandarin Chinese test itself, in addition to not having stimuli representative of the /i/ category for L1 Mandarin Chinese, make up issues for this method. As a result, it proved difficult to compare the Mandarin Chinese and Norwegian results at the front acoustic space (see Chapter 5). By the time this misjudgment became clear, it was much too late to redo the tests.

2.3.1 The procedure

Before undergoing manipulation, the vowels were first isolated from the original sentences using Audacity. To do a resynthesis of a vowel in Praat, the following procedure is followed (see appendix 4 for command details. File names are given in parentheses below for easy identification in the appendix): A studio recording of a vowel is synthesized into a new sampling frequency of 10 000 Hz (vowel_resampled). A problem that frequently occurred during this process was significant fluctuation in the formant values. For one re-sampling, the alteration to the F2 was as much as 400 Hz. This was mostly due to unstable formants that crossed other formants in small “drops”. I assume that Praat then overcompensated by lowering the formant value of the affected formant, F2, to avoid collision of formants F2 and F3. In the figure below we see the formant grid of a Mandarin Chinese /y/:

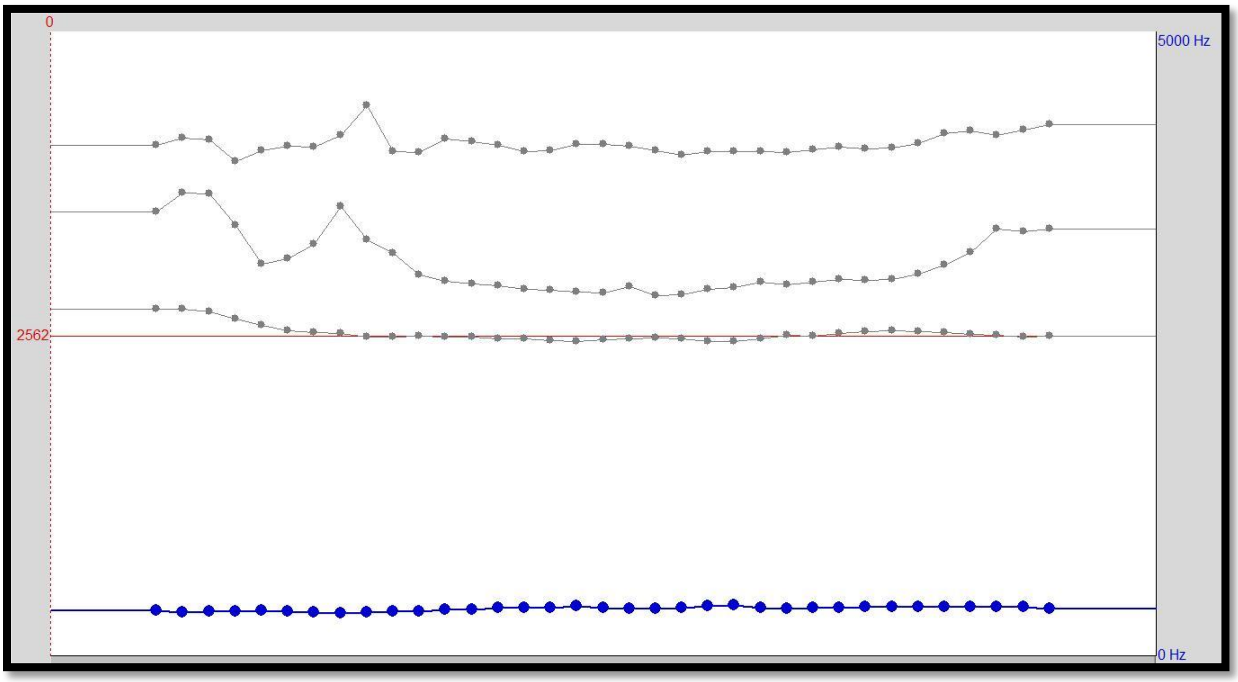

Fig 2.3 A Mandarin Chinese /y/ before re-sampling in Praat.

As we can see, there are some unstable points here, but these points do not cross other formants. In the next picture we see the re-sampling of the same sound. Praat does a good job evening out the formants here, creating a more stable environment for synthesis.

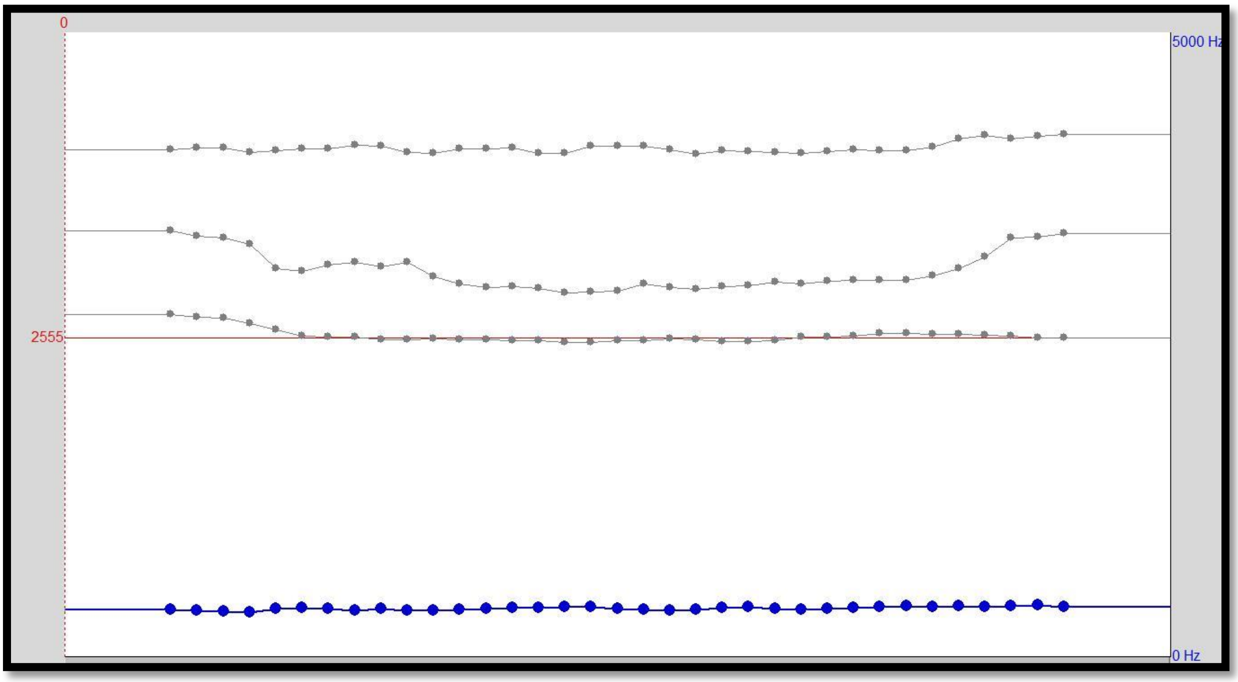

Fig 2.4 A Mandarin Chinese /y/ after re-sampling in Praat

For this vowel, the most deviant sections at the beginning and end were removed from the signal, creating a stable vowel ready for synthesis. This new sound file is then run through a linear-prediction analysis with a prediction order of 10 (vowel_LPC). The resampled vowel and the LPC vowel are then combined and the result is our source sound for resynthesis (vowel_source).

The next step is to manipulate the formant values of the vowel. The resampled sound is analyzed to formant (vowel_resampled_formant), before the new file is converted to a formantgrid (vowel_resampled_formantgrid). As mentioned earlier, formant manipulation is an automatic process in Praat where you plot in how much you want each value to change, and then Praat carries out that command. The F2 values between segments were altered in sequences of 100 Hz. This amounted to 19 stimuli for the Mandarin Chinese test, and 20 for the Norwegian. Praat extracts the average value of a selected area, here the entire vowel, and then alters the formants' average values. Praat was not always consistent in its alterations, and sometimes the formants were altered by only 50 Hz or 120 Hz despite the command of altering 100 Hz. It was the average value of the vowel that had to be manipulated 100 Hz per stimuli, so I made small corrections to the commands, for instance by manipulating 150 Hz for some stimuli, but only 50 for others. The average values of the finished stimuli ended up being in steps of approximately 100 Hz (but see section 2.3.1.1**)**.

Finally, the altered formantgrid and the source file are filtered, creating the final file (Vowel_Resampled_Source_Filt), ready for placement in the frame sentence. The dB of the vowel has changed through manipulation, and must therefore be adjusted in Audacity to agree with the frame sentence. Any gaps between sound waves in the transitional area are repaired, and the stimulus is ready.

The same vowel should be used as a base for the entire continuum, in order to avoid any other acoustic traits affecting the samples (e.g. duration). One would want to have the vowel with the highest formant values as the base vowel. Starting with a vowel with low formant values, and then increasing the F2 value gradually, could cause F2 and F3 to ”crash” during manipulation. This in turn would cause erroneous manipulation as Praat would be unable to separate the two formants, and thus not be able to alter the correct formant. Therefore, /y/ seemed to be the best starting point for both manipulations.

2.3.1.1 Mandarin Chinese stimuli

For Mandarin Chinese, the first attempt of manipulation of /y/ failed because the recorded vowels were corrupted in one way or the other. For example, one of them had a significant “dive” in F2 in the middle of the vowel, making it unfit for manipulation. An attempt to manually stabilize the vowel in the formant grid was made, but that ruined the sample, creating a metallic sound to it. New recordings were made, and several vowels tested for manipulation. One /y/ was found suitable, and manipulation went without problem from that point on. I ended up with the following results:

The original values of the Mandarin Chinese /y/:

F2: 2585

F3: 3148

The values after resampling and cutting some deviant areas at the beginning and end:

F2: 2542

F3: 3002

As we can see, the differences between the original and the resample are small, and thereby this resample is suitable for further manipulation. The resampled /y/ was thus used as the starting point for further manipulation.

There were some problems in that the input command did not match the output values. Consequently, some adjustments had to be made in the input. The stimuli with the file names CY4, CY6, CY7, CY8, CY11 were all altered by only -50 Hz for F2, while CY9, CY10 were altered by -150 Hz. CY3X is a “forced” stimulus because it was specifically manipulated from the resample to get that exact value. Creating an interval with exactly 100 Hz interval proved impossible, as no degree of fine tuning or adjustments gave me the exact wanted results. Some intervals were under 100 Hz, others over, as seen in table 2.1 below:

F2	Average value
Resample	2542
CY1	2547
CY2	2385
CY3	2259
CY3X	2172
CY4	2074
CY5	1936
CY6	1861
CY7	1703
CY8	1600
CY9	1497
CY10	1412
CY11	1329
CY12	1262
CY13	1188
CY14	1033
CY15	984
CY16	814
CY17	758
CY18	631

Table 2.1 The Mandarin Chinese Stimuli

The continuum /y/ - /u/ is as follows: CY1 → CY18, resulting in 19 stimuli (CY3X included).

2.3.1.2 The Norwegian stimuli

The creation of the Norwegian stimuli proved a little more difficult. Several /y/ samples were tried for resynthesis, but none of them turned out optimal. As discussed above, /u/ was not a suitable base for manipulation towards /y/, and one would want the same base vowel throughout the continuum. Thus, /ʉ/ was chosen to be resynthesized both towards /y/ and towards /u/. The only prior modification of this vowel was deletion of an unstable part at the end of the vowel. The resample values of this vowel were acceptable:

The original values of /ʉ/:

F2=2180

F3=2822

The resampled values of /ʉ/:

F2: 2214

F3=2774

The resampled /ʉ/ was used as a starting point for further manipulation. In the continuum from /ʉ/ to /u/, 15 intervals were created, but the stimulus with the file name NU5 was deleted because the output from Praat was a stimulus that was too close to neighboring stimuli. The modification for stimuli NU9 and NU10 were -50 Hz, while the rest were manipulated with −100 Hz for F2. The result is shown below:

F2	Average value
NU0	2122
NU1	2053
NU2	1956
NU3	1899
NU4	1793
(NU5)	(1724)
NU6	1652
NU7	1518
NU8	1402
NU9	1324
NU10	1237
NU11	1120
NU12	1012
NU13	899
NU14	781

Table 2.2 The Norwegian stimuli, set 1

For the continuum between /ʉ/ and /y/, which also had the starting point of the resampled /ʉ/, the result was eight stimuli. Stimulus NY3 was deleted because of close proximity to the surrounding stimuli. Stimuli NY1 to NY3 were manipulated with +50 Hz, while the remaining stimuli underwent a change of +100.

F2	Middelverdi
NY0	2214
NY1	2335
NY2	2433
(NY3)	(2491)
NY4	2518
NY5	2586
NY6	2650
NY7	2709
NY8	2754

Table 2.3 The Norwegian stimuli, set 2

The continuum /y/ - /ʉ/ - /u/ is as follows: NY8 → NY0 → NU0 → NU14.

2.4 The perception test

The perception tests were also created in Praat. For this purpose, I used a template script provided by Professor van Dommelen at NTNU (see appendix 5 for the full script). The options were orthographic, something that possibly made correct identification more difficult (see section 1.1.2). [20] As the words used were nonsense words, using images was not an option.

The native Mandarin Chinese test consisted of the options ‘ny’ and ‘no’ for 95 stimuli. The native Norwegian test consisted of the options: ‘ny’, ‘nu’ or ‘no’ for 100 stimuli. The Mandarin Chinese version of this test also included the option ‘ni’. Flege (2003a) points out that “too many labels might bias subjects away from the correct answer, while too few labels can lead to underreporting” (p. 23) The choice to include ‘ni’ for the Mandarin Chinese version of the Norwegian test was to avoid underreporting, but as we will see in chapter 3, it seems like it lead to bias instead.

[20] See section 3.2 on how it was attempted to circumvent the orthographic issues.

A series of stimuli (see section 2.3) were presented in random order to the informants in a self-paced, labeling experiment (Ashby & Maidment 2005: 181). For each acoustic input, they were asked for a forced judgment of ‘Which word are you hearing now?’. They were instructed to guess in cases where they were uncertain, and to take as much time as they needed. They could listen to each sentence twice if they so chose, and they were given a self-paced break for every 10 stimuli. They made their choices by clicking on the options on a computer monitor (see picture 2.5 below)

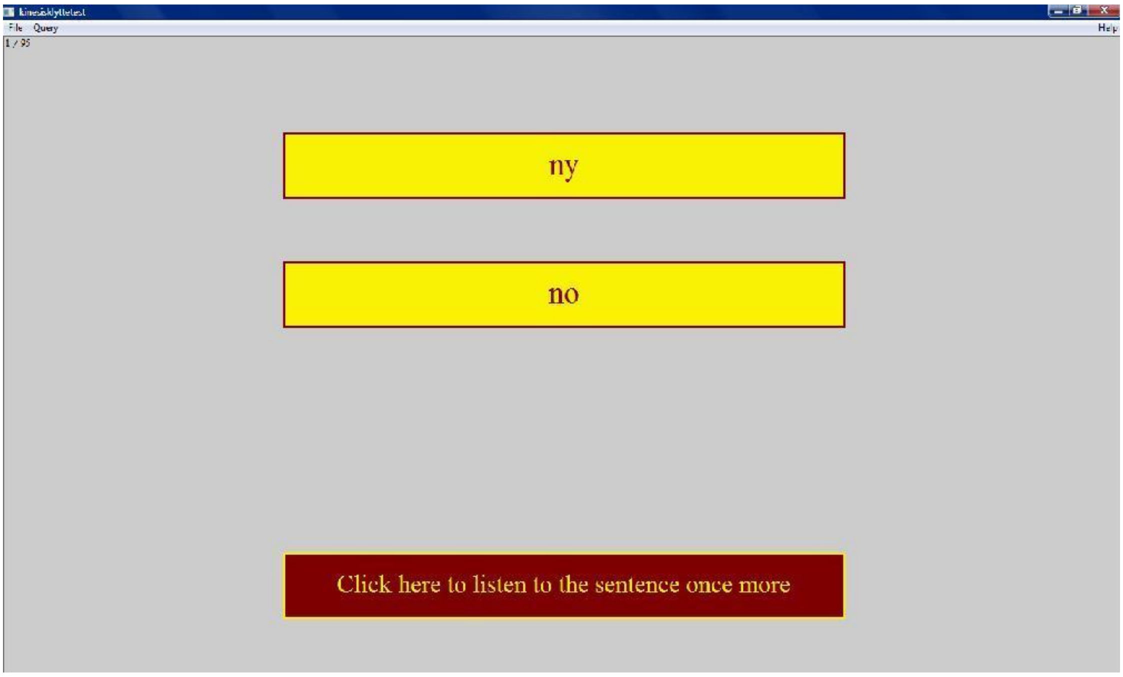

Fig. 2.5 A screenshot of the Mandarin Chinese perception test.

Before each test, the original sentences were played for the informants and they were simultaneously shown in writing what orthographic letter each target vowel corresponded to. This was done to avoid confusion and misunderstanding concerning the difference between sounds and orthography (ref. Chapter 1).

In retrospect, a training round prior to the actual test could have been beneficial. Here, the informants could have become familiar with the mechanics of the test, and given the option to ask clarifying questions before performing the actual test. The informal production test which was performed prior to the testing could have been done once more between the training and the test. The unfamiliar situation could be the reason for the great fluctuation seen in the early rounds of the tests for some informants. Furthermore, an additional production test could have ensured further that the informants were able to successfully pair sounds and orthographic representation. It appeared that even native categories were sometimes not successfully established. Lack of ability to identify the correct pairing seems the most reasonable explanation for this result.

Furthermore, adding more stimuli could have enhanced the results. A greater amount of data is expected to have yielded more stable results, and thus facilitate a more substantiated analysis. Having only 5 stimuli per category was however a conscious choice. These informants were not compensated economically for their time, and the test was repetitive. The latter factor is suspected to cause decreased concentration if the test included a large amount of stimuli. After conducting the tests, this seemed to be the correct choice as some informants did express boredom. Some informants also went from stable choices in the first 3-4 rounds, to suddenly give unstable responses in the last rounds. When conducting similar tests in the future, doubling the stimuli seems preferable, and then one may extract the answers given in the middle section. This is because there was some confusion in the early rounds, and then some choices in the later rounds that seemed “forced”, i.e. ‘ni’ suddenly being chosen in the last round.

Another option is doing ABX tests instead. Here, the informants listen to stimuli in sets of triangles: Play one manipulated stimuli (A), then another manipulated stimuli (B), before playing the prototype of the target sound (X). The informants are then asked whether A or B resemble X more. This method of testing is shown to produce good results for studies on category boundaries, e.g. Best et.al (2003). The choice of not applying this method was based on the same considerations as those outlined above: the amount of time it would take for the unpaid informants.

2.5 Informants

For this study, a total of 17 informants were needed, and the initial selection was systematic: one person from each language to record the stimuli, 5 native Norwegians formed a control group and 10 native speakers of Mandarin Chinese, 5 with little training in Norwegian and 5 with substantial (3-4 courses) training in Norwegian. The final count was 5 Norwegians, and 6 Mandarin Chinese with different levels of training. Such a small informant base makes the present research a qualitative study rather than a quantitative one.

2.5.1 Recordings

The initial stage of the study, where recordings of the test sentences were carried out, required one native speaker of Mandarin Chinese and one native speaker of Norwegian. As formant frequencies vary between male and female speakers (Ashby and Maidment 2005: 72), all of these recordings had to be of the same gender to avoid notable differences in Hertz values. The choice fell on female speakers as I decided to use myself as the Norwegian informant for the recordings. By doing so, I could alter the recording method if necessary before recording the Mandarin Chinese sentences. I am from the western part of Norway, but 'normalized' my pronunciation to be as near Urban Eastern Norwegian pronunciation as possible during the recordings to avoid dialectal influence on the perception tests. The 'normalization' included avoidance of dialect forms and an attempt to resemble the Eastern Norwegian intonation as closely as possible.

The Chinese female speaker is a fellow master student who volunteered to participate in this study. She is from a large city in the eastern part of China and had resided in Norway for 15 months at the time of the recording. Keeping in mind that China is a vast country with many dialectal differences (see section 3.2) the choice of informants for the perception tests would ideally have to be based on a similar geographical location as this informant to get accurate results. Her geographical origin also excludes her formant values from comparison with the cited literature in chapter 1.1.2, which is Beijing Mandarin.

2.5.2 Perception tests

The native Norwegian test was the first to be carried out. A group of 5 native Norwegians were needed to establish category boundaries for the Norwegian segment and to act as a control group. These informants were recruited locally, that is, my own Norwegian friends and family. NM_1 commented that he found the volume of the test a bit low, and it was therefore adjusted. None of the other informants had any questions or concerns during the tests.

The Mandarin Chinese perception tests required 5 speakers from both level 1 and level 3 of the course 'Norwegian for Foreigners'.[21] Gender did not matter in the selection as perception is not dependent on this factor. The informants were recruited in January, February and March 2013 after a new round of Norwegian courses had started. Finding informants proved a difficult task. A first attempt to recruit informants was done via the teachers of the Norwegian course at the end of January. This resulted in one informant. For the second attempt, over 30 e-mails were sent to students of all course levels in February. This yielded one more informant. The third and final attempt was made in March, when the Chinese woman who

[21] http://www.ntnu.edu/norwegiancourse

recorded the Mandarin sentences contacted the student group for Chinese students at NTNU. At this point, the criteria for informants were now simply “Mandarin Chinese natives either trained or untrained in Norwegian”, and this resulted in the final four informants. As for the levels of training, there was some variation. Three of the informants were at level 1, one at level 2, one at level 3 and one at level 4.

Chapter 3 Experiment and results

This study is concerned with initial ranking and following reranking of constraints pertaining second language learning. Ideally, such research should follow the same informants throughout their gradual learning of the second language in a diachronic study. However, as the time frame of a master's thesis is relatively short, the choice fell on a synchronic apparent time study instead. This means choosing different groups of informants based on similar backgrounds, but with different levels of L2 competence. There are several shortcomings to this choice, especially seeing that the informant base ended up being quite small. This makes it difficult to make any generalizations surrounding unlearned vs. learned informants, and we are reduced to focus on individual differences. However, some patterns emerged even within this small informant base. Furthermore, this study being apparent time only gives us the option to hypothesize how an informant has advanced in learning steps, especially when it comes to the learned informants.

Describing absolute formant values for all users of one language is an impossible task. One thing to consider is individual differences, and even differences for each individual from one time to another. Then we have the added factor of dialects. Both Norwegian and Mandarin have large dialectal differences, but to what degree these differences affect vowel perception is to the best of my knowledge unknown. In any case, as an exchange student in Trondheim, a city with a great influx of students from all over the country, the likelihood of being exposed to all Norwegian dialects is quite high.

From the results presented below, we see that the Norwegian /u/ category is quite small in the southern part of Norway, while it is a bit bigger further north. However, my results are hardly representative seeing there were only five informants, one from each area. For Mandarin, most of the informants in the present study categorize the Mandarin Chinese /u/ up until 1400-1500 Hz, but the informant from Bengbu (see map below) has a smaller /u/ category at 1033 Hz, and the informant from Chongqing has a large /u/ category ending at ca. 1703 Hz. They are all geographically linked to different dialects (see the sections for each informant below), but as there is only one informant per dialect, no hypotheses are possible regarding dialectal differences.

Fig.3.1The geographical locations of all the Mandarin Chinese informants, http://goo.gl/maps/47KPQ, retrieved 03.04.13.

With a large amount of informants, one can make generalizations through finding the average values. Lee and Zee (2001) chose to add standard deviations in addition to mean values (section 1). The present thesis does not have a large informant base, and will not have sufficient data to introduce standard deviations.

Consequently, the focus will lie on individual results and the results will be interpreted as follows: For overall responses, categories are deemed clear/ideal if the responses to a stimulus are in a 2/3 favor of one vowel, and that this trend continues for subsequent stimuli. For the individual responses, 4/5 is deemed a clear category, while 3/5 is unsettled. For example, in the overall Norwegian results (see section 3.1), the stimulus 1120 is categorized as /ʉ/ because that was the response in 17 out of 25 stimuli. Contrarily, in the overall Mandarin Chinese results (see section 3.2), where the responses to stimulus 1497 are 16 for /u/ and 14 for /ʉ/, the stimulus is deemed "unsettled". It might be a transitional acoustic area, where it is difficult to categorize the stimuli, or the informant is unable to establish a category. There are more instances like this throughout the Mandarin Chinese results and it is evident that these areas cannot be deemed "clear" or "ideal".

3.1 Norwegian native category boundaries

Before the perception test, the 5 Norwegian informants were given oral instructions of what sound constituted what orthographic vowel. The Norwegian informants were not given the choice of ‘ni’ due to this option being added after these tests had been carried out. An extra native Norwegian test that included this last option was done to see if it had a large effect on the results. This Norwegian informant answered ‘ni’ only once, and it was at the very end of the test. The possibility that the Mandarin Chinese informants would not identify the stimuli as /i/ if the option was not there remains an open question.

The overall responses for the Norwegian native test were as follows:

Norwegian	Number of responses		
F2 Hz value	/u/	/ʉ/	/y/
781	25		
899	25		
1012	15	10	
1120	8	17	
1237	9	16	
1324	6	19	
1402	4	21	
1518	2	23	
1652		25	
1793		25	
1899		25	
1956		25	
2053		24	1
2122		19	6
2214		1	24
2335			25
2433			25
2518			25
2650			25
2754			25

Table 3.1 The native Norwegian results

According to the overall Norwegian results, the /u/ category is quite small. It is clear up until 899 Hz, then quickly declines until it is not perceived anymore at 1652 Hz. The stimulus 1012 is an “unsettled” area. The Norwegians’ /ʉ/ category, on the other hand, is quite large, spanning from 1120 Hz to 2122 Hz, with an additional 4 perceptions at 1012 and 1 perception at 2214. /y/ is starting to be perceived at 2053, and dominates from 2214. From this, we conclude that the ideal category for the Norwegian /u/ is up until 899, the ideal /ʉ/ category spans from 1120 – 2122, and the ideal /y/ category starts at 2214.

In the figure below, the results are plotted into a visual continuum, from lowest to highest frequency. The grey/shaded area denotes an unsettled area, i.e. not 2/3 in favor of one vowel. In the visual figures, all values are presented as a round number, i.e. 781 Hz is presented as 800 Hz and 2214 Hz as 2200 Hz.[22] These round values are what will be employed in the discussions in chapters 4 and 5.

			/u/		/ʉ/	/y/
L1 NO			800 - 900	1000	1100 - 2100	2200 - 2700

Fig 3.2 The native Norwegian category boundaries

The values of the stimuli for the Norwegian continuum are not equal to those of the Mandarin Chinese (e.g. there is a stimulus 814 Hz in Mandarin Chinese, and a stimulus 899 Hz in Norwegian). This is a problem that will be discussed in Chapter 6, and the background for the uneven stimuli is explained in Chapter 2.

3.1.1 Individual variations

As we will be looking into individual perceptions in Mandarin Chinese, a brief look at individual differences in Norwegian is necessary. The full numbers will not be shown here, but they can be viewed in appendix 6.

NM1: From Bergen (West). He has a small category for /u/, dominating the area from 781 to 899 Hz. He then has an overlapping area at 1012 Hz. The /ʉ/ category is large, spanning from 1120 to 2053 Hz. There is one confused area at 2122 Hz. Then /y/ takes over from 2214 Hz.

NM1		/u/		/ʉ/		/y/

Fig 3.3 The native Norwegian category boundaries for NM1

[22] There were some problematic values, such as the continuum 1956 Hz - 2053 Hz, where they were both rounded to 2000 Hz to achieve a continuum with all values from 800 Hz to 2700 Hz present.

NM2: From Fauske (North). NM2 has a large /u/ category, spanning from 781 to 1324 Hz. He has an overlapping area between 1402 and 1518 Hz. /ʉ/ then dominates the area from 1652 to 2122 Hz. /y/ is then perceived from 2214 Hz outwards.

Fig 3.4 The native Norwegian category boundaries for NM2

NM3: Fredrikstad (East). Again, we see a small category for /u/, covering the area from 781 to 899. /ʉ/ is very large, spanning over the area from 1012 to 2122. /y/ then takes over.

Fig 3.5 The native Norwegian category boundaries for NM3

NM4: Ålesund (North West). There is a small /u/ category here as well, from 781 to 899. /ʉ/ covers the area of 1012 to 2122. /y/ covers the end of the continuum.

Fig 3.6 The native Norwegian category boundaries for NM4

NM5: Trondheim (Mid). This informant shows the most overlap between categories. /u/ dominates from 781 to 1120 Hz, but is perceived as high up the continuum as 1402. There is an area of overlapping identifications between 1237 and 1324 Hz. /ʉ/ dominates from 1402 to 2053, but is perceived from 1120 to 2122. /y/ is perceived from 2053, but dominates first from 2122.[23]

Fig 3.7 The native Norwegian category boundaries for NM5

From the individual results we conclude that /u/ should not be perceived after 1324 and /y/ should take over at 2214. For the analysis in chapter 5 we will rely on the overall results, but the dialectal differences will be discussed in section 5.4.

[23] Due to an error, this informant did the Norwegian perception test intended for the Chinese Mandarin informants. This test includes the option <ni>. NM5 did, however, only answer <ni> once, and this answer is for the sake of analysis omitted in the results. It is presented in appendix 6.

3.2 Mandarin Chinese category boundaries

I played the original recordings in both languages prior to each test and pointed to the orthographic representations the different vowels constituted as they listened to the vowels. After doing the test with four informants, it seemed like these instructions did not suffice as the results varied greatly. The only informant who seemingly got it "right" (in the L1 test) was a linguistics student (CF3), and thus familiar with the difference between phonetic symbols and orthography. For the remaining two tests, the informants were therefore instructed to repeat the sounds to me so that it was certain it was clear what sound each orthographic vowel represented. The initial tests were not redone as it was impossible to gather more informants, and the informants who had already participated would have been biased in their choices following a discussion about the test after the first completion. The extra instructions did not, however, seem to result in more stable answers.

From informant 3 and onwards, I also instructed them that not all of the choices presented to them were necessarily present. They were told that the perception of the different vowels could vary from individual to individual, and that this was OK. I chose this approach because I had done a perception test with a German student[24] who pointed out that she started to worry half-way through that she had not clicked <ni> yet, so she started clicking <ni> from that point on. Furthermore, both CF2 and CM6 commented after the test that they relied on visual cues to distinguish between Norwegian /i/ and /y/ in "real life" situations, indicating that some of their focus had been on the i/y difference. However, recognizing the close front vowels as either /y/ or /i/ does not affect this study as it is the category boundary of the /ʉ/ category that is important. One could also say that we are looking for the category boundaries between close front, close central and close back vowels, where both /i/ and /y/ belong to the super-category 'front'.

Informant CF2 also seemed preoccupied with the fact that the Mandarin Chinese words were nonsense words. She laughed a couple of times doing the test and as a result I told her that the Norwegian stimuli were not nonsense words but meant "nine"- /ni:/, "new" - /ny:/, "now" (in old form) - /nʉ:/, and "now" (in new form/dialect) - /nu:/. I kept doing this for the other informants as well, explaining that finding existing minimal pairs in Mandarin Chinese with these exact vowels in that exact consonant environment that also carried the same tone was impossible. They all agreed with me in that.

[24] Initially, this thesis was supposed to include both German and Mandarin Chinese, but due to lack of informants German had to be dropped from the study.

Lastly, one of the informants, CF1, asked me: "I'm hearing the same sound all the time, is this normal?" during the first round. I explained that it was possible she heard similar sounds many times after one another as it was a random selection. I told all the informants this from that point on. CF1 also seemed visibly unsettled by me staying in the room during the tests, so I moved to an adjacent room, having the door open, for the subsequent tests.

The analysis in this study is primarily concerned with individual rankings in both L1 and L2, and additionally the differences between learned and unlearned Mandarin Chinese students of Norwegian. However, a clear pattern has emerged in the overall results, and is thus dedicated some space in both this presentation and subsequent discussions.

Mandarin Chinese		
F2	/u/	/y/
631	29	1
758	30	
814	28	2
984	28	2
1033	29	1
1188	27	3
1262	24	6
1329	23	7
1412	23	7
1497	16	14
1600	7	23
1703	6	23
1861	4	26
1936	2	28
2074	1	29
2172	3	27
2259	2	28
2385	1	29
2547		30

Norwegian				
F2	/u/	/ʉ/	/y/	/i/
781	30			
899	26	4		
1012	27	3		
1120	28	2		
1237	26	4		
1324	27	3		
1402	27	3		
1518	25	5		
1652	19	11		
1793	13	17		
1899	11	16	3	
1956	3	22	5	
2053	1	25	4	
2122	3	22	5	
2214	1	25	4	
2335		16	13	1
2433		14	15	1
2518	1	12	14	3
2650		12	11	7
2754		12	12	6

Tables 3.2 & 3.3 The overall native responses to the Mandarin Chinese perception test (left), and the Mandarin Chinese responses to the Norwegian perception test (right).

From this we can see that in the case of their native language, the Mandarin Chinese informants favor /u/ until the area around 1497, and after that recognize /y/ considerably more than /u/. Looking at clear categories, where the criteria is 20 identifications out of 30 (2/3, as discussed above), the native /u/ ends at 1412, and /y/ starts at 1600, with 1497 being unsettled.

For the Norwegian vowels, /u/ has significant identification by the Mandarin Chinese until around 1652 Hz, but the clear /u/ category ends at 1518 Hz. /ʉ/ is favored in the area from 1793 to 2335, but is a clear category only from 1956 to 2214 Hz. After that, there are small differences in the perception of /ʉ/ or /y/. There are also some occurrences of /i/, but perception of /i/ never dominates. There is a small unsettled area between /u/ and /ʉ/ from 1652 – 1793, before some /y/ perception starts at 1899. The Mandarin Chinese fail to establish a clear category for /y/, and the stimuli from 2335 and up are all unsettled.

Figure 3.8 below show these results plotted into a visual continuum. As in the figures in section 3.1 above, shaded areas denote unsettled categories (here not 20/30 responses in favor of one category) and all values presented in the figure are round numbers. One square equals one stimulus (approximately 100 Hz). Since the Norwegian continuum, contrary to the Mandarin Chinese continuum, does not include a stimulus at 600 and 700 Hz (see section 2.1), the Mandarin Chinese continuum starts after two squares. Likewise, the Mandarin Chinese continuum did not include stimuli at 2600 and 2700 Hz. The Mandarin Chinese also has some intermediate values for their stimuli, leaving the continuum with one less stimulus than the Norwegian continuum (see Chapter 2). These intermediate values are more prominent in the upper part of the continuum, thus the Mandarin Chinese continuum ends three squares (three stimuli) prior to the Norwegian.

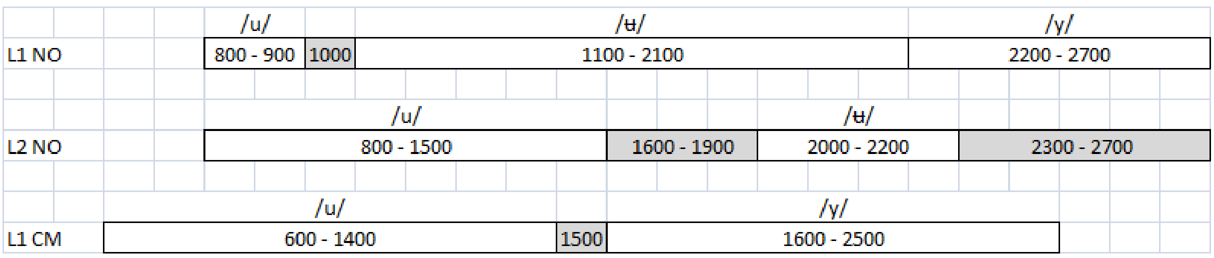

Fig 3.8 The overall results of the Norwegian perception of Norwegian category boundaries (L1 NO, top), the Mandarin Chinese perception of the Norwegian category boundaries (L2 NO, middle), and the Mandarin Chinese perception of Mandarin Chinese category boundaries (L1 CM, bottom)

The similarity between how the Mandarin Chinese perceive their native /u/ category and their perception of the Norwegian /u/ category is substantial, indicating a degree of assimilation.[25] It is clear that the informants are aware of the /ʉ/ category, and have established a category for it. The end part of the Norwegian continuum resembles what Flege (2003a) calls response at chance level. Where it was expected that the /ʉ/ category might not have been established, the results rather indicate that it is the Norwegian /y/ category that has not been established.

3.2.1 Informant CF1 – Level 1

This informant is 23 years old and from Qin Huangdao, in the Hebei province. This area has the Northeastern Mandarin dialect, which is close to Standard Chinese. She has studied English before. Her native language is Mandarin, and she speaks decent English, but she did not write that on her questionnaire. She has lived in Norway for 6 months, and never lived here before. She is not surrounded by Norwegian at all at a daily basis, and says that this is because all her classes are taught in English.

CF1 was insecure doing the test, asking many questions along the way. She also seemed quite bored throughout the test, sighing loudly as she clicked.

[25] The notion of assimilation in second language acquisition will be presented and discussed in Chapter 4.

Mandarin Chinese		
F2	/u/	/y/
631	4	1
758	5	
814	4	1
984	4	1
1033	5	
1188	5	
1262	5	
1329	5	
1412	5	
1497	3	2
1600	1	4
1703		5
1861	1	4
1936		5
2074		5
2172		5
2259		5
2385		5
2547		5

Norwegian				
F2	/u/	/ʉ/	/y/	/i/
781	5			
899	2	3		
1012	3	2		
1120	4	1		
1237	4	1		
1324	3	2		
1402	4	1		
1518	3	2		
1652	1	4		
1793	2	3		
1899	1	4		
1956		5		
2053		4	1	
2122		4	1	
2214		5		
2335		2	3	
2433		1	4	
2518		1	3	1
2650			1	4
2754			1	4

Tables 3.4 & 3.5 The native responses to the Mandarin Chinese perception test (left), and the Mandarin Chinese responses to the Norwegian perception test (right) by informant CF1.

In her native language, she has a clear /u/ category up until 1412, where an unsettled area at 1497 follows. The /y/ category starts at 1600 and reaches almost perfect identification. These results show that she understood the mechanisms of the test, but she also had some severe misperceptions, with /y/ at 631 Hz being the most notable. Interestingly, she is a perfect match to the overall results that we saw in section 4.2 above.

Her Norwegian responses fluctuate greatly, thus clear categories are difficult to extract from her responses. Her /ʉ/ identifications are at chance level, fluctuating throughout most of the continuum. The only clear, consistent area of a /ʉ/ category is between 1899 and 2214. Following the previous outlined principles of categorization, the /u/ category only spans from

1120- 1237. The /y/ category is also not clear, the only dominant response being at 2433. /i/, however, dominate the last section of the continuum, from 2650 – 2754. A visual figure of her continuum is given in figure 3.9 below:

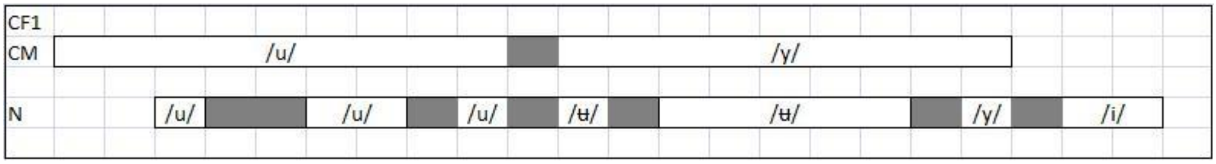

Fig 3.9 CF1's category boundaries in Mandarin Chinese (CM) and Norwegian (N)[26]

Looking at the order in which she gave the responses, we see that she had a clear /u/ category in the first two rounds, spanning from 781 – 1899 the first round, then to 1518 the second, before /ʉ/ appears increasingly frequently in the last three rounds. Contrarily, the /y/ category is not established until the third round, not being perceived at all in the first round, with the most /y/ responses at the last round.

Despite there being more fluctuation in the Norwegian /u/ category, we can see clear similarities to the native /u/ category, with last perception at around 1800 Hz, and domination ending around 1400. Her perception of /ʉ/ spanned from 899 – 2518, but she only perceived /ʉ/ as the dominant vowel consistently from 1899 to 2214. It seems that CF1 felt that she *should* perceive /ʉ/, either from the test or her knowledge about Norwegian, or both. This knowledge seems to have overshadowed almost all other phonological knowledge, preventing her from creating stable categories.

3.2.2 Informant CF2 – Level 4

This informant is 31 years old and from Chongqing. This area has the Upper Yangtze Mandarin dialect, and it has some differences from the standard Mandarin dialect. She has studied English, and speaks Chinese, English and Norwegian. She has lived in Norway for 2 years, and also lived in Norway for 10 months in 2008. She says that she is sometimes surrounded by Norwegian during a normal day. This informant spent a great deal of time on the test and seemed interested in doing well.

[26] In this and future figures of individual results, Hertz values will not be added. The reader is referred to the tables for numbers, and reminded that one square equals approximately 100 Hz.

Mandarin Chinese		
F2	/u/	/y/
631	5	
758	5	
814	4	1
984	4	1
1033	5	
1188	5	
1262	5	
1329	5	
1412	5	
1497	5	
1600	4	1
1703	4	1
1861	3	2
1936	2	3
2074	1	4
2172	3	2
2259	2	3
2385	1	4
2547		5

Norwegian				
F2	/u/	/ʉ/	/y/	/i/
781	5			
899	5			
1012	5			
1120	5			
1237	5			
1324	5			
1402	5			
1518	5			
1652	4	1		
1793	4	1		
1899	4	1		
1956	1	4		
2053	1	4		
2122	2	2	1	
2214	1	4		
2335		2	2	1
2433			4	1
2518			3	2
2650		1	2	2
2754			4	1

Tables 3.6 & 3.7 The native responses to the Mandarin Chinese perception test (left), and the Mandarin Chinese responses to the Norwegian perception test (right) by informants CF2.

In her native language, she clearly favors /u/ up until 1703 Hz, followed by a confused area from 1861 – 2259 Hz, before /y/ dominates for the last two. Looking at the order of her responses, no clear pattern emerges. There are severe misperceptions at both the first and last rounds, so it is likely that she experienced problems with matching the correct sound to the correct alphabetical representation.

In Norwegian, CF2 clearly favors /u/ up until 1899, and has thus established a category for the Norwegian /u/. It is worth noting that /u/ is perceived as far up as 2214, but seldom. The categories of /ʉ/ and /y/ cannot be said to have been established. /ʉ/ is only just perceived from 1652 to 1899, and dominates from 1956 to 2053, and again at 2214 Hz. The value of

2335 is confused, with /ʉ/, /y/ and /i/ all perceived. /ʉ/ is also perceived once at 2650. /y/ only dominates at 2433 and 2754 Hz. /i/ is perceived once or twice from 2335 – 2754.

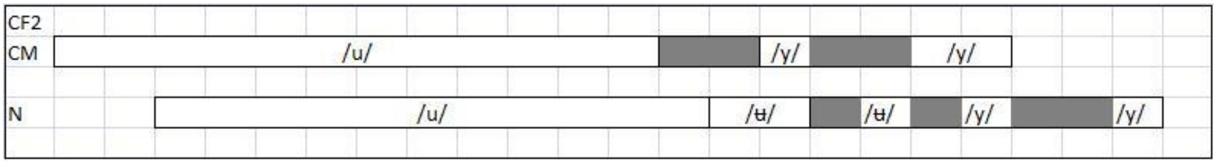

Fig 3.10 CF2's category boundaries in Mandarin Chinese (CM) and Norwegian (N)

We see that the /u/ category is quite similar in both languages for this informant. Both her perceived ideal categories and last perceptions of /u/ end at around the same areas in both languages. Furthermore, the higher values are unsettled in both languages.

3.2.3 Informant CF3 – Level 2

CF3 is 26 years old and from Bengbu, in the Anhui province. This area has the Zhongyuan Mandarin dialect. She has studied Japanese, and speaks Mandarin and English. She has lived in Norway for 2 years, and not lived in Norway before this. She says she speaks a little to her roommate in Norwegian; otherwise there is no Norwegian input in her daily life.

This informant was left alone in the room after been given instructions because the previous two informants seemed self-conscious when I was present. I checked on her regularly and also came in to start the Norwegian test after she finished the Mandarin Chinese one. She had no further questions and seemed to carry out the test well.

Mandarin Chinese		
F2	/u/	/y/
631	5	
758	5	
814	5	
984	5	
1033	5	
1188	3	2
1262	1	4
1329		5
1412		5
1497		5
1600		5
1703		5
1861		5
1936		5
2074		5
2172		5
2259		5
2385		5
2547		5

Norwegian				
F2	/u/	/ʉ/	/y/	/i/
781	5			
899	5			
1012	5			
1120	5			
1237	4	1		
1324	4	1		
1402	4	1		
1518	2	3		
1652	3	2		
1793	1	4		
1899		2	3	
1956		2	3	
2053		3	2	
2122		2	3	
2214		3	2	
2335		2	3	
2433		1	4	
2518	1	1	3	
2650			5	
2754			5	

Tables 3.8 & 3.9 The native responses to the Mandarin Chinese perception test (left), and the Mandarin Chinese responses to the Norwegian perception test (right) by informants CF3.

This informant has clear categories in Chinese. She exclusively perceives /u/ from 631 – 1033. Then, there is confusion at 1188, before she stops perceiving /u/ after 1262. /y/ is first perceived at the unsettled value of 1188, and dominates from there on out. This indicates that she understood the mechanisms of the test, and also successfully paired sound signal and alphabetical representation.

For Norwegian, her categories are not as clear. The only clear category of some size is /u/, from 781 to 1402 Hz. She starts perceiving /ʉ/ from 1237, and she stops perceiving /u/ after 1793 (but there is one identification at 2518). There is an unsettled area from 1518 to 1652, and then /ʉ/ dominates at 1793 Hz. After that, she is largely unable to hear the difference between /ʉ/ and /y/. /y/ only dominates at 2650 and 2754. There are no /i/ occurrences.

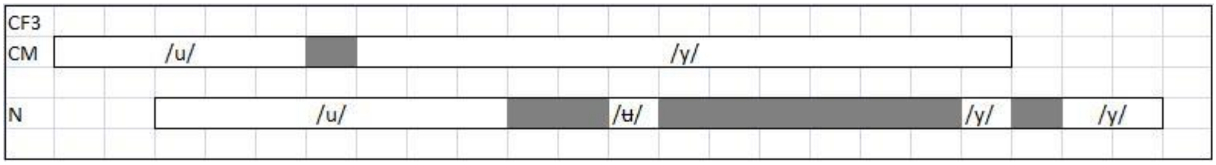

Fig 3.11 CF3's category boundaries in Mandarin Chinese (CM) and Norwegian (N)

CF3 starts to perceive the Norwegian /ʉ/ in the area where she set the boundary between /u/ and /y/ in her native language, but only just. At the higher end of the continuum, her results are overall at chance level, indicating that she is having problems distinguishing between /ʉ/ and /y/. That /u/ has such a large category might be due to assimilation (see Chapter 4).

3.2.4 Informant CF4 – level 3

This informant is 34 years old and from Xinjiang, Shihezi. This area has both the Lan-Yin dialect and the Zhongyuan dialect. She has studied English, Dutch and Norwegian, and speaks English and Mandarin. She has lived in Norway for 2 years, and lived in Norway for 6 months at an earlier time. She says she is surrounded by Norwegian to a small degree on a normal day. CF4 also mentioned that distinguishing between /i/ and /y/ was hardest for her. I sat in an adjacent room during the tests.

Mandarin Chinese		
F2	/u/	/y/
631	5	
758	5	
814	5	
984	5	
1033	5	
1188	4	1
1262	4	1
1329	4	1
1412	5	
1497	2	3
1600	2	3
1703	2	3
1861		5
1936		5
2074		5
2172		5
2259		5
2385		5
2547		5

Norwegian				
F2	/u/	/ʉ/	/y/	/i/
781	5			
899	5			
1012	5			
1120	5			
1237	5			
1324	5			
1402	5			
1518	5			
1652	3	2		
1793	1	4		
1899		5		
1956	1	4		
2053		5		
2122		5		
2214		5		
2335		4	1	
2433		3	2	
2518		1	4	
2650			5	
2754			5	

Tables 3.10 & 3.11 The native responses to the Mandarin Chinese perception test (left), and the Mandarin Chinese responses to the Norwegian perception test (right) by informants CF4.

In her native language, CF4 has a clear /u/ category from 631 to 1412, but with a couple of /y/ perceptions in the area. This is followed by a confused area from 1497 – 1703. /y/ dominates slightly, however. From 1861, /y/ is the only perceived vowel.

In Norwegian, CF4 has a clear /u/ category from 781 to 1518. There is some confusion at 1652, before /ʉ/ takes over as the clear category. /u/ is still slightly perceived up until 1956. She has established a category for /ʉ/ from 1793 – 2335, followed by an unsettled area at 2433, before the /y/ category dominates the remainder of the continuum.

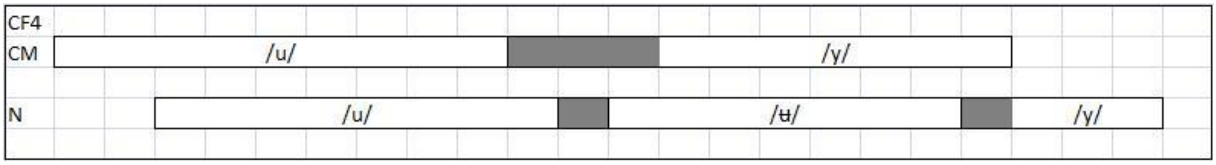

Fig 3.12 A visual figure of CF4's category boundaries in Mandarin Chinese (CM) and Norwegian (N)

The category boundary between the native /u/ and /y/ bear resemblance to that of her Norwegian boundary between /u/ and /ʉ/, and can be credited to /u/ assimilation. /ʉ/ has taken over the acoustic space of the Mandarin Chinese /y/, with the Norwegian /y/ only being categorized above that of the, tested, Mandarin Chinese /y/.

3.2.5 Informant CF5 – level 1.

This informant is 28 years old and from Shanghai. This area is described for the Wu language, but the informant says she is a native speaker of Mandarin. She has studied English, and speaks Mandarin and English. She has lived in Norway for about 4 years, and has not lived in Norway before this. She states that she is surrounded by Norwegian "a little" during a normal day. She did not attend NTNU's Norwegian course, but a similar course at Folkeuniversitetet, which lasted for four months.

I sat in an adjacent room during the tests. CF5 did not perceive a single Norwegian front vowel. Interestingly, I forgot to tell this informant that all stimuli might not be present. Furthermore, she also mentioned that she had problems with the difference between /y/ and /i/, indicating that she sees the Norwegian /y/ as far more front than it is. As discussed in Chapter 2, there were no /i/ stimuli in the perception tests.

Mandarin Chinese		
F2	/u/	/y/
631	5	
758	5	
814	5	
984	5	
1033	4	1
1188	5	
1262	4	1
1329	4	1
1412	3	2
1497	2	3
1600		5
1703		5
1861		5
1936		5
2074		5
2172		5
2259		5
2385		5
2547		5

Norwegian				
F2	/u/	/ʉ/	/y/	/i/
781	5			
899	5			
1012	5			
1120	5			
1237	5			
1324	5			
1402	5			
1518	5			
1652	5			
1793	4	1		
1899	5			
1956	1	4		
2053		5		
2122	1	4		
2214		5		
2335		5		
2433		5		
2518		5		
2650		5		
2754		5		

Tables 3.12 & 3.13 The native responses to the Mandarin Chinese perception test (left), and the Mandarin Chinese responses to the Norwegian perception test (right) by informants CF5.

In CF5's native language, /u/ clearly dominates from 631 to 1329, and is then followed by a confused area from 1412 – 1497. After that, /u/ is no longer perceived. /y/ is perceived as far back as 1033, but very seldom. The responses indicate that she to some degree correctly paired sound signal and alphabetical representation.

For Norwegian, this informant has very clear categories, but none of those are front. /u/ spans from 781 – 1899, being last perceived at 2122. /ʉ/ is perceived from 1793, and dominates from 1956.

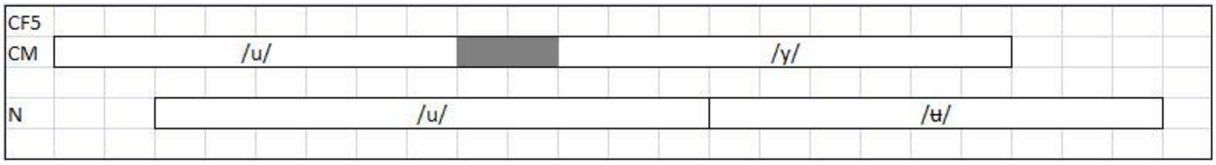

Fig 3.13 A visual figure of CF5's category boundaries in Mandarin Chinese (CM) and Norwegian (N)

One hypothesis about these responses is that the informant has the knowledge about /ʉ/'s presence in the Norwegian language, and thus chooses this over /y/. It is possible that she finds the Norwegian /y/ different from the Mandarin Chinese /y/ to such a degree that she substitutes the /y/ perception with /ʉ/. Whether this is due to acoustic signals or the alphabetical representation is unknown.

3.2.6 Informant CM6 – level 1

This was the first and only Chinese male informant. He is 25 years old and from Daqing, in the Heilongjigang Province, in the North-East part of China. This area has the Northeastern Mandarin dialect, which is very close to Standard Chinese. He has studied languages, but did not state which. He speaks Mandarin and English and has lived in Norway for 1 ½ years. He has not lived in Norway before. He states to be surrounded by Norwegian to a small degree on a normal day.

He was eager to do the test and seemed interested. He asked to hear the original sentences twice to make sure "his Norwegian hat was on". This informant also commented on the difficulty of distinguishing between /y/ and /i/.

Mandarin Chinese		
F2	/u/	/y/
631	5	
758	5	
814	5	
984	5	
1033	5	
1188	5	
1262	5	
1329	5	
1412	5	
1497	4	1
1600		5
1703		5
1861		5
1936		5
2074		5
2172		5
2259		5
2385		5
2547		5

Norwegian				
F2	/u/	/ʉ/	/y/	/i/
781	5			
899	4	1		
1012	4	1		
1120	4	1		
1237	3	2		
1324	5			
1402	4	1		
1518	5			
1652	3	2		
1793	1	4		
1899	1	4		
1956		3	2	
2053		4	1	
2122		5		
2214		3	2	
2335		1	4	
2433		4	1	
2518		4	1	
2650		1	3	1
2754		2	2	1

Tables 3.14 & 3.15 The native responses to the Mandarin Chinese perception test (left), and the Mandarin Chinese responses to the Norwegian perception test (right) by informants CM6.

CM6 has very clear categories for his native language. /u/ spans from 631 – 1497, where /y/ is perceived once at 1497. /u/ is then no longer perceived, and /y/ takes over.

The Norwegian boundaries are not so clear, and there is some confusion throughout the continuum. With the exception of the 1237 stimulus, the /u/ category spans from 781 to 1518 Hz. From 1652 and throughout, there is great variation, with /ʉ/ being the mostly dominant response, but with /y/ interruptions and unsettled areas throughout.

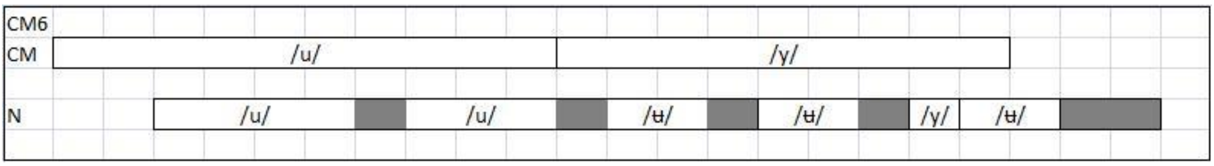

Fig 3.14 A visual figure of CM6's category boundaries in Mandarin Chinese (CM) and Norwegian (N)

We can here see a degree of assimilation, as the native /u/ ends around the area where domination of Norwegian /u/ ends. As we have seen with other informants, /ʉ/ also here dominates the acoustic space of the native /y/.

3.3 Summary

None of the Mandarin Chinese informants, regardless of level, has correctly established all three Norwegian close vowel category boundaries. Most of them have in common that the Norwegian /u/ category dominates the continuum, and is often expanded beyond that of the native /u/ category. This is in contrast to how the native Norwegians categorized a small /u/ category. Furthermore, the Norwegian /y/ is a weak category for the Mandarin Chinese. It is the least perceived overall and in some cases (CF4) not perceived at all. All informants have, to different degrees, established a category for the unknown category /ʉ/, as seen in figure 3.15 below, where the shaded parts represent areas that do not hold clear categorization.

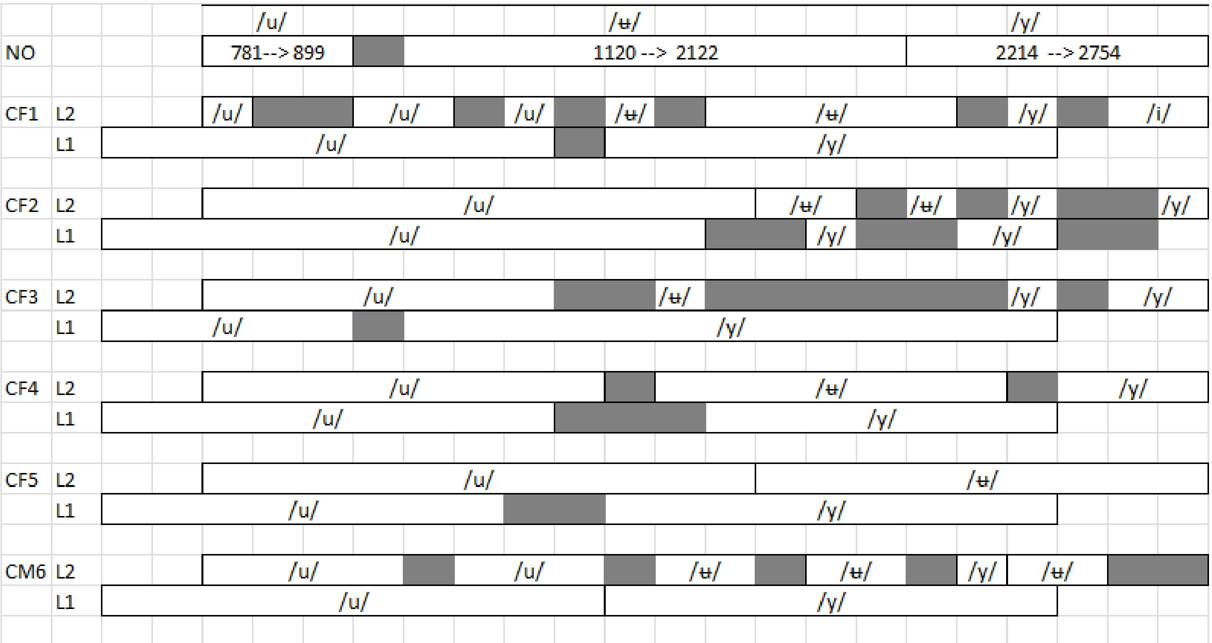

Fig 3.15 The overall native Norwegian results (NO), compared with the L1 (Mandarin Chinese) and L2 (Norwegian) results of the individual informants

A feasible conclusion from these numbers is that the Norwegian /u/ category, together with the lower value section of the /ʉ/ category, has undergone assimilation to the Mandarin Chinese /u/ category for the Mandarin Chinese informants. The Norwegian /y/ category, however, seems to have been dissimilated from the Mandarin Chinese /y/ category, and in many instances been replaced with the new /ʉ/ category. These results will be discussed in light of theory in the next chapter.

Chapter 4 Theoretical discussion

The first part of this chapter will look at perceptual skills in general, and outline some existing theories and hypotheses on perception in the L2. In the second part, the theories used for the analysis in Chapter 6 will be presented. Some of the discussions in this chapter are based on the results from the experiments presented in chapter 3.

4.1 Perception and learning of L2

There are many factors affecting second language acquisition, some of which researchers agree on, and some where they differ. One thing most researchers concur on is that the L1 influences the L2 during L2 acquisition, and prominent models within this view are the L2LP (Escudero 2005), PAM (Best 1994, Best et al. 2001), SLM (Flege 1995, 2003a, 2003b, 2007) and NLM (Kuhl et al. 1991, 1995, 2000, 2003). These models will be outlined in this section, together with a general overview of some other factors that may affect second language perceptual skills.

4.1.1 Perceptual Skills

When a learner first encounters the L2, he is on level 1 of the listening process and can only receive the sound signals through hearing and is not able to categorize the sounds according to the L2 system (Husby & Kløve 1998: Ch. 7). As the learner becomes more skilled in the language, he moves up to level 2 and is able to perceive the sound signals and consequently analyze and categorize the sounds. The perceptual skills can be affected by both inner and outer factors, as well as the general knowledge about the L2 and its culture.

Physiological, psychological, social and linguistic factors all shape how a learner is able to acquire a second language (Husby & Kløve 1998: Ch.1). Within physiological factors we find critical age, motor skills and neurology. The neurological area of plasticity is something that we will encounter in the Gradual Learning Algorithm as what decides the learning rate (see section 4.2.1). As the brain gets older, it loses plasticity, meaning that it gets harder to create new neural pathways. These pathways enable us to learn new things, like languages. Escudero (2005) suggests that as long as the amount of L2 input exceeds the lack of plasticity, learning will happen.

Socio-psychological factors can play as big a role as physiological. How a learner mentally meets the L2 can have great effect. If the learner identifies with the language and is positive to the culture of L2, there will be affective reasons to acquire the language. He may attach his own sense of ego to his proficiency of the L2, and this will motivate him to further learning as not to embarrass himself in a conversation in the L2.

Unlike production, perception is not monitored by the environment and will therefore be subjected to social control. A misperception will thus continue unchanged unless it is mirrored in the production, or the learner himself becomes aware of the error due to sufficient L2 exposure. Social factors overall are key to successfully acquire a new language, notably time in the L2 culture, to what degree the learner is exposed to the L2, how much the learner uses his L1 in the L2 acquisition process and the duration of formal instruction in the L2. From the questions asked in the questionnaire (appendix 2), we see that the informants lack L2 input in their everyday lives, and this may be the reason why not even the level 3 students have established clear Norwegian categories.

Another possible confusing element is orthography. The Chinese writing system is Hanzi, logograms, while the Norwegian writing system is Latin, an alphabet. In Pinyin, the official system of writing Chinese characters using the Latin alphabet, the Mandarin /y/ is mostly written as <ü>. Contrarily, in Norwegian the same symbol, "y", is used for phonetic transcription and for Norwegian orthographic writing.

4.1.2 Existing theories and hypotheses on perception of L2

Mandarin Chinese has two rounded close vowels, while Norwegian has three. In production, the front-rounded vowel, /y/, and the back rounded vowel, /u/, are nearly identical in Norwegian and Mandarin Chinese (see Chapter 1). This indicates that it is the central rounded category, /ʉ/, which is the unknown category that has to be established.[27] From the perception results in Chapter 3, however, it is clear that the Mandarin Chinese informants are experiencing trouble in the acquisition of the Norwegian /y/ category. This goes to show that it is the contrast between a front rounded category and a central rounded category that has to be acquired.

Consequently, there are primarily two tasks for the acquisition of the Norwegian close rounded vowels for the Mandarin Chinese: Shift L1 category boundaries for the front and back rounded vowel categories to mimic those of the L2, and integrate the new L2 contrasts in an already categorized dimension. The optimal outcome of this learning situation is that an L1 category is split to make room for the new contrast between the front rounded /y/ and the

[27] The reader is at this point reminded that the phonetic labels, /y ʉ u/ will be used as short forms for the phonological categories front rounded , central rounded and back rounded, as was discussed in Chapter 1.

central rounded category, /ʉ/. This learning step requires the new category to be established in the close acoustic space, and this new category is argued to be /ʉ/ based on production values (see Chapter 1).

The tasks are thus to shorten the acoustic space of the L1 categories, and allot the remaining acoustic space to the new category, /ʉ/. Furthermore, the L1 categories which are phonetically similar to the L2 categories, /y u/, should be stored in the L2 phonology as *new* categories particular to Norwegian, which adhere to the category boundaries of the L2.

Both of these tasks can result in assimilation instead. Assimilation in the L2 sense is that properties of the L1 categories are transferred onto the L2 categories because the learner is unable to distinguish between them, (Flege, 2003a). The L2 category is thus *equated* with the L1 category. Consequently, one or more of the L2 categories will not be established. This is most likely with similar categories, i.e. categories that are acoustically close. In the present study, the new L2 category /ʉ/ resides in the middle of the acoustic space of L1 /y/ and /u/, with the median value of the /ʉ/ category being 1600 Hz, and the category boundary between L1 /y/ and /u/ being at approximately 1500 Hz. The Norwegian /ʉ/ is thus acoustically close to two L1 categories. This is illustrated in figure 4.1 below,[28] which is based on data from Chapter 3:

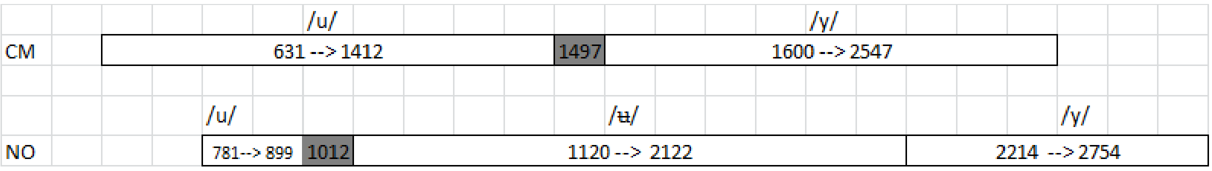

Fig 4.1 The boundaries of Mandarin Chinese native categories (CM) and the boundaries of Norwegian native categories (NO) as given by the overall results of the respective languages' native speakers.

For a person who was not aware of Norwegian having an additional category, we would expect that values from 1500 Hz and upwards would assimilate to L1 /y/, and values from 1500 Hz and downwards would assimilate to L1 /u/. This assimilation is suggested to be a result of learners' inclination to perceive a new language through their L1 phonological and phonetic filter. The most prominent theories within this view are presented below.

Kuhl et al. (1991, 1995, 2000, 2003) found that sounds in the L2 that are similar to sounds in the L1, i.e. perceptually close, will be assimilated in the L2 learning sense. She attributes this

[28] The formalisms behind the figure are explained in section 3.2

to the magnet effect of native-language phonetic prototypes to which these similar sounds are drawn towards, and named the model the Native Language Magnet Model. By studying infants, adults and monkeys, Kuhl (1991) discovered that humans exhibit a perceptual magnet effect on phonetic prototypes when categorizing speech categories. When infants learn their mother tongue, they create perceptual maps where the acoustic dimension is warped towards their native language's prototypes (2000). Thus, the infants "develop perceptual and cognitive processes that are specialized for their language", (2003: B48). When learning a new language, then, the child not only perceives this language through a native filter where the magnet effect alters the perceived distances, but the child will also rely on native cues and may not pick up on the cues exploited in the new language (1995, 2003).

Flege's Speech Learning Model (SLM; 1995, 2003a, 2003b, 2007) works under the assumption that the L1 and L2 subsystems are simultaneously active and may interfere with one another.[29] The L2 sounds are perceived through the L1 filter, and are therefore initially perceptually assimilated by an L1 category. Flege's studies have shown that learners may also filter out features (cues) that are not needed in the L1, if they at all have access to these features. As the learner acquires a larger L2 vocabulary, however, L1 filtering will diminish as the need for differentiating increases (2003b: 9).

Escudero (2005), unlike Flege, assumes that the learner has two separate systems for L1 and L2, and that the L1 remains stable through L2 development. However, the interlanguage state during initial exposure is a result of both systems being activated simultaneously (p. 114f). When the learner is in an absolute beginner stage, Escudero posits that the learner goes through a stage of full copying where all L2 sounds are mapped to L1 sounds (2005: 98). This copying will gradually decrease as the learner gains negative evidence, i.e. semantic learning.

According to Flege (2003), the establishment of a new category is more likely if there is some perceptual distance between the L1 and L2 sounds (p. 10). However, if the learner fails to discern between the sounds, "phonetic category assimilation" may happen instead where a new "composite" category is created, combining the features of both the L1 and the L2 sound (2003b:12). This seems to be the case of the Norwegian L2 /y/ and /ʉ/ categories for many of the Mandarin Chinese informants in the present study (see figure 4.2 below). Their replies to the values in the acoustic space of the Norwegian /y/ show they are largely unable to

[29] Flege also discusses the critical age and whether or not such a thing exists. This area of acquisition will not be explored in this thesis.

distinguish between /y/ and /ʉ/, ending up with an acoustic space that is a bit of both categories. This is illustrated in figure 3.15, reproduced as figure 4.2 below. The figure is based on data from Chapter 3, where the results were discussed in more detail.

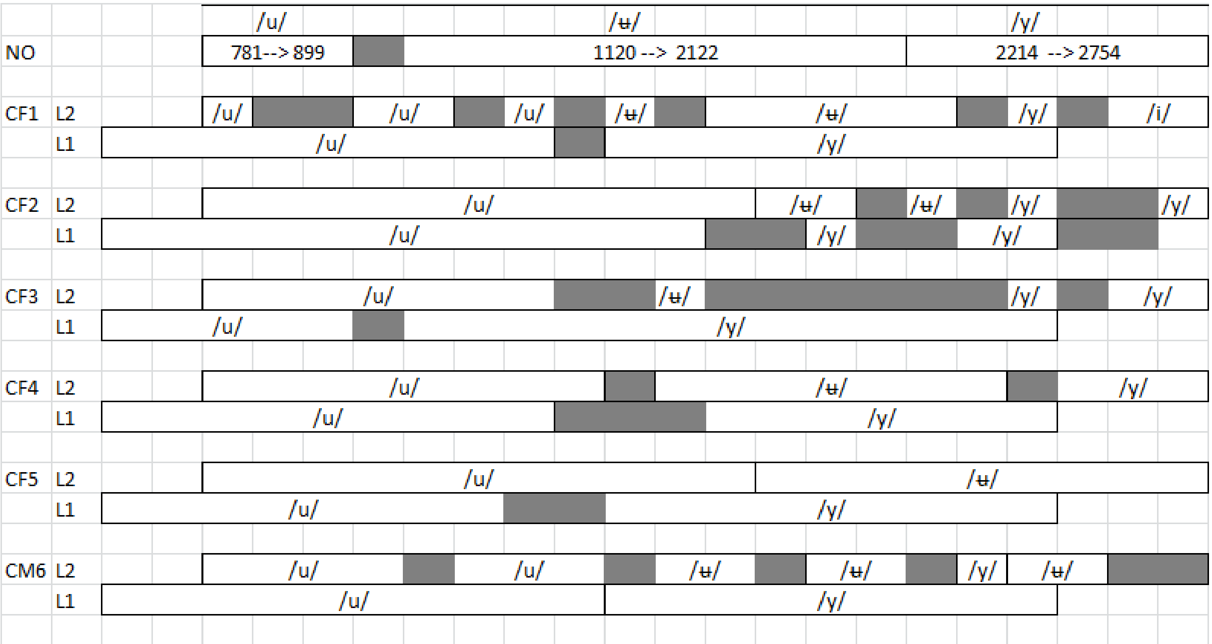

Fig 4.2 The overall native Norwegian results (NO), compared with the L1 (Mandarin Chinese) and L2 (Norwegian) results of the individual informants

Kuhl (1995, 2003) suggests that the degree of difficulty in creating new categories in the L2 depends on the proximity of that category to an L1 magnet. The outcome of L1 and L2 categories that are perceptually close may be assimilation (Kuhl 1991). As seen in figure 4.2 above, the result for the Mandarin Chinese informants in the present study seems to be assimilation of the lowest values of the L2 /ʉ/ category to the L1 /u/ category, as the stimuli replies for L2 /u/ largely mirror the replies to the stimuli for L1 /u/. This indicates that L1 /u/ acts as a magnet on the lowest values of the L2 /ʉ/. Additionally, large parts of the L1 /y/ category are identified with the L2 /ʉ/ category. The /ʉ/ identification in the L2 (and L1) acoustic space of /y/ cannot be explained by the magnet effect to an L1 category as /ʉ/ is an L2 category, but can rather be explained as conscious knowledge overriding the phonological knowledge (see Chapter 5).

The Perceptual Assimilation Model (PAM) posited by Best (1994), looks at how well infants discern between native and nonnative contrastive phones. If a listener perceives the sounds as similar to a native category, they will be assimilated, i.e. be perceived as equivalent. If the sounds are not familiar, the listener will identify discrepancies and thus not assimilate the sounds to native categories. Best (1994) proposes four patterns of assimilation: The Two Categories type, where the nonnative sounds are assimilated to two native categories, the Single Category type, where two nonnative sounds are assimilated to one native sound, the Category Goodness type, where both nonnative sounds are categorized to one native category, but one nonnative sound is deemed a better fit to that category than the other nonnative sound, and the Non-Assimilable type, where the nonnative sounds are too different from native sounds to be assimilated.

In the present study, the contrastive pair /u/-/ʉ/ falls within the Single Category type, as both L2 /ʉ/ and /u/ values assimilate to L1 /u/ (see figure 3.15, repeated above as figure 4.2 above). However, this is only for the lower Hertz values of /ʉ/, and not the entire category. The higher part of the /ʉ/ category seemingly belongs to the Non-Assimilable category as it is not assimilated to any native category. Rather, the native category /y/ is assimilated to the L2 /ʉ/. Interestingly, this resembles Single Category assimilation, but in this case it is assimilation to a nonnative sound and not a native sound. Such occurrences are not included in PAM, but are in the present study accredited conscious knowledge about the nonnative category (see discussion is Chapter 5).

In her dissertation, *Linguistic Perception and Second Language Acquisition,* Escudero (2005) looks at how new cues help establish a new category in the L2 and proposes a Second Language Linguistic Perception model, the L2LP. She implements this model in the Gradual Learning Algorithm (see section 4.2.1). She identifies three scenarios, the SIMILAR scenario (same categories in L1 and L2), the SUBSET scenario (less categories in the L2 than the L1) and the NEW scenario (more categories in the L2 than the L1). Escudero (2005) predicts the following tasks and degrees of difficulty for the NEW and SIMILAR scenario:

L2LP proposal	Prediction for NEW	Prediction for SIMILAR
Initial state	Too few categories	Same number of categories
Perceptual task	1. *Create* perceptual mappings 2. *Integrate* auditory cues	Adjust category boundaries
Representational task	1. *Create* phonetic categories 2. *Create* segments	None
Degree of difficulty	Very difficult	Not difficult

Fig 4.3 Comparative initial states and learning tasks in the NEW and SIMILAR scenarios (Escudero 2005: 258)

In this study, we encounter what Escudero (2005) calls the NEW scenario (p.155 ff): The Mandarin learners have fewer vowels in their L1 than what they encounter in the L2, Norwegian. The learner will start out with phonemic equation, where the unknown L2 category will be identified with an existing L1 category. According to figure 4.3, the learner is faced with both a perceptual task, defining category boundaries, and a representational task, creating new categories and segments.

All the informants in the present study have, to varying degrees of success, created a new category for /ʉ/, but how? Escudero (2005) offers no explanation for this in her thesis, as she predicts that "the L2 development and the L2 state in a NEW scenario are restricted to cases that involve at least one non previously-categorized auditory dimension, such as vowel duration in Spanish learners of SBE [Standard British English] vowels." (Escudero 2005: 161). She states in her conclusion that explaining splitting of categories is beyond the scope of her study (p. 317).

Vowel duration cannot be seen as an auditory dimension responsible for the creation of the new categories in the present thesis. This is because the Norwegian close-central categories, the categories that are not in the Mandarin Chinese vowel inventory, are both long and short. Consequently, the Mandarin Chinese will have to acquire a category that does not differ in cues in the L1 and L2, namely the short /ʉ/. The cues for this category are F1, F2 and F3, of which F2 is tested for in this thesis (see section 2.1). It is a possibility that the Mandarin Chinese rely on F3 more than F2, or a combination of the two, while categorizing the close categories, but as discussed in section 2.1, the differences between categories in terms of F3 are small in both languages. However, an experiment with F3 or a combination of F2 and F3 might yield a different result than what is seen in the present thesis.

Lastly, the amount of input of L2 language is crucial in both Escudero and Flege, thus this was one of the questions the informants were asked in the present study's questionnaire. Common for all of them is that they are surrounded by Norwegian to a lesser degree in a normal day. This might lead to what Escudero calls 'fossilization', a state of learning where the process stops before mastering native-like understanding (2005: 114). This possibility can only be identified if a learner is followed through stages of learning, which is not feasible within the frames of this thesis.

4.1.3 Earlier research on L2 perception of Norwegian vowels

Best et al. (2003) studied how native speakers of English, French and Danish categorized and discriminated the Norwegian pairs /i/-/y/, /y/-/u/, /y/-/ʉ/,/ʉ/-/u/. None of the subjects had any training in Norwegian. The aim of the study was to investigate the phonological and phonetic effects of the listener's native language when perceiving nonnative vowels. The subjects were tested through ABX discrimination tests.[30] Best et al. (2003) found that the Americans, who only use /i/ and /u̟/ of the close vowels, assimilated L2 /y/ to L1 /i/ and L2 /ʉ/ to L1 /u̟/. The Danish natives, who use the close vowels /i y u/, assimilated L2 /ʉ/ to L1 /y/, and in 53 % of the cases assimilated L2 /u/ to L1 /o/. The French language also uses /i y u/, and the native French subjects assimilated the L2 /ʉ/ to the L1 /y/. The L2 /y/, however, was assimilated to the L1 /i/.

The French results are the most interesting to the present study, as it resembles Mandarin Chinese both in the phonological systems of /i y u/ and that the phonetic realizations of /y/ are not as protruded as the Norwegian (and Danish). The results from the present study are not directly comparable to the results from the study done by Best et al., however, as the informants in the latter study were not trained in Norwegian. Importantly, all /ʉ/ stimuli in Best et al. were assimilated to a native category, and none identified as the L2 /ʉ/. As we saw in chapter 3, the L2 /ʉ/ stimuli of the present study were rarely identified as belonging to a /y/

[30] In an ABX test, the informant is played a triad of stimuli, A and B which are manipulated vowels, and X which is the "prototype". The informant is then asked which of A and B resemble X more.

category.[31] They were, however, often identified as belonging to a /u/ category, which resembles the results of the Americans in Best et al. Still, it is important to note that the stimuli in the present study were in a continuum, and the L2 /u/ identifications were in the lower part of the continuum, near L1 /u/, while the stimuli in Best et al. were one stimulus per optimized vowel.

In another study, Albertsen (2008) looked at how native speakers of French, Spanish and German perceived vowel length contrast and lip-rounding contrast in Norwegian vowels. When testing for their ability to distinguish between rounded and unrounded close vowels, he found that the native Spanish listeners had a significantly lower amount of correct answers than the other groups. He also found that in the continuum test, the Norwegian and German listeners showed a good ability to distinguish between phonological short and long vowels, while the Spanish and French could not.

4.2. Optimality Theory

The presentation in this section assumes that the reader is familiar with Optimality Theory. The examples used throughout this chapter are the overall results of the Mandarin Chinese and Norwegian informants from the experiment in the present study (see Chapter 3).

Proposed by Prince and Smolensky in 1993, Optimality Theory (OT) is a model which claims that the difference between languages is one of ranking of universal constraints. In this model, the input, or underlying form, is processed by GEN(erator) which generates output candidates which are evaluated by EVAL(uator) according to the language-specific ranking of the violable constraints, CON(straints). CON is divided into markedness and faithfulness constraints, where markedness constraints have been identified by linguistic universality – patterns that occur systematically in languages. For instance, the central close vowel /ʉ/ is less common in languages, so there could theoretically exist a markedness constraint *ʉ (/ʉ/ is not allowed, but see discussion about Flemming's Dispersion Theory below, and the discussion about phonological categories, here central rounded, in Chapter 1.). All constraints exist in any and all grammars, even if they are presumably inactive.

[31] In this comparison, it is not relevant to which language, L1 or L2, the categories belong to, as the information we are looking for is to what category the new category maps to, and not where the category boundaries are.

The candidate with the least severe violations, i.e. the candidate that satisfies the topmost constraints the best, is considered the optimal candidate and emerges as the output, or surface form. In tableau formalism, violations are marked with an asterisk, *, and fatal violations are marked with exclamation points, !. The winner is marked with ☞, and an incorrect winner marked with 💣. When we have an incorrect winner, we use ✓ to mark the candidate that should have won.

The tableau below illustrates the basic mechanisms of OT, as presented above. The illustration is based on data from chapter 3, and the analysis is in accordance with Boersma and colleagues (see section 4.2.1 below). This type of analysis differs slightly from classical OT in that the output is not a phonetic realization, but a perceived vowel category. It is important to note that the illustration below is a perception grammar and not a production grammar, as one would normally expect from an OT analysis. The input, 1200 Hz, is the second formant value that the learner hears, and this analysis shows a Mandarin Chinese informant perceiving 1200 Hz as Norwegian /u/,[32] while he should have perceived a Norwegian /ʉ/.[33]

1200 Hz	[1200 Hz] not /ʉ/	[1200 Hz] not /u/
✓ /ʉ/	*!	
💣/u/		*

Tableau 4.1 An incorrect perception of the Norwegian value 1200 Hz

Tableau 4.1 illustrates the ranking: ‘[1200 Hz] not /ʉ/’ >> ‘[1200 Hz] not /u/’, where >> stands for “higher ranked than”. If we find no proof for ranking ‘[1200 Hz] not /ʉ/’ over ‘[1200 Hz] not /u/’, i.e. if we get the same optimal candidate regardless of the ranking of those constraints, the ranking would not be determined: ‘[1200 Hz] not /ʉ/’ , ‘[1200 Hz] not /u/’, where the comma stands for “equally ranked”. This would be formalized in the tableau by a dotted line between said constraints instead of a solid one as in the tableau above.

[32] The Chinese Mandarin overall results showed 26 identifications of /u/ and 4 identifications of /ʉ/ at 1237 Hz.

[33] The Norwegian overall results showed 16 identifications of /ʉ/ and 9 identifications of /u/ at 1237 Hz.

4.2.1 Gradual Learning Algorithm (GLA):

Boersma and colleagues (1997, 1999, 2001, 2003, 2004) developed the Gradual Learning Algorithm (GLA) to explain the learning process of languages within OT. GLA is based on reranking of constraints when there is a mismatch between the learner's output and the adult (correct) output. The knowledge of such mismatch is given to the learner by perceptual learning, and only by identifying that there is an error, will the learner be able to rerank the constraints. When such an error is detected, the learner will demote violated constraints whose violation predicts the wrong winner candidate, and promote the constraints that do not violate the winning candidate, shown by the arrows:

[1200 Hz]	[1200 Hz] not /ʉ/	[1200 Hz] not /u/
💣 /u/		←*
✓ /ʉ/	→*!	

Tableau 4.2 A learning step in the perception grammar

The constraints used here are *cue constraints*. The cue constraints are of the form '[number Hz] not /category/', where /category/ is e.g. /close back rounded/, but abbreviated /u/. There is one cue constraint per hertz-category pair. For example, if we have a cue constraint '[1200 Hz] not /u/', all candidates that are not /u/ when the input is 1200 Hz will be penalized.

Based on the learning step above, the learner's ranking is altered, and the output is a new correct winning candidate:

[1200 Hz]	[1200 Hz] not /u/	[1200 Hz] not /ʉ/
/u/	*!	
☞ /ʉ/		*

Tableau 4.3 A complete learning step

This symmetric promotion/demotion happens in small steps decided by plasticity (see section 4.1.1 above). It is important to note that the Gradual Learning Algorithm assumes several learning cycles to complete learning processes. For example, in Weiand (2007), 180 000 training repetitions were carried out in the Praat scripting language (p. 9). Out of those repetitions, learning took place in only 6 percent of the trials. In the present study, only current rankings are shown, (see Chapter 2 about apparent time).

4.2.1.1 Stochastic Optimality Theory

GLA assumes a Stochastic version of Optimality Theory (Boersma & Hayes 2001). This version explains variation, optionality and probability within a language. Stochastic Optimality Theory (SOT) is an important component for GLA to work in a computational environment, but not for the static analysis presented in this thesis. Still, the basic mechanics of SOT is presented below because this version offers a framework of analysis for the "undeterminable" areas in the Mandarin Chinese perception of the acoustic area of the L2 Norwegian /ʉ y/.

In the traditional OT, all constraints have in principle clearly delimited rankings, e.g. C1 >> C2 >> C3, or in case of constraints with no internal ranking: C1, C2, C3. Applying the continuous ranking scale, however, gives us a scenario where some constraint can be higher ranked to a higher degree than the others. Consider figure 4.4 below:

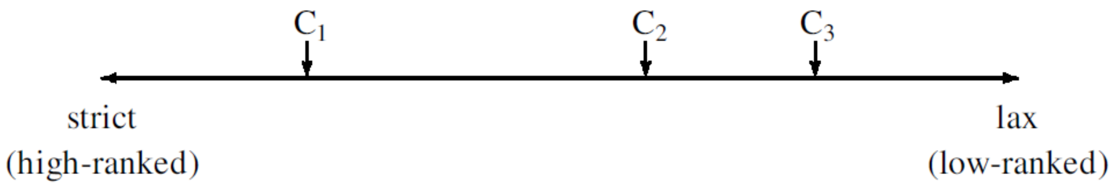

Fig 4.4. Categorical ranking of constraints (C) along a continuous scale (Boersma & Hayes 2001: 47).

Here, C1 has a higher ranking value than C2 and C3, thus C1 outranks C2 to a higher degree than C2 outranks C3. This will not have much effect unless we assign ranges of values to the constraints and these ranges overlap.

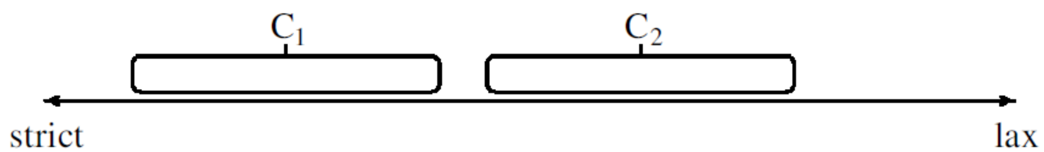

Fig 4.5. Categorical ranking with values (Boersma & Hayes 2001: 47)

Boersma & Hayes (2001) decide these ranges by Gaussian distribution where there is a single, most probable, peak in the center, and then a gentle but swift decline in probability towards zero in both directions of the curve. Real world noisy events (mumbling, general noise, etc.) are taken into account and a *hypothetical* value of how much it affects the listeners' perception is set per evaluation. This value is related to the evaluation itself, not the individual constraints and is added to the computer simulation. Boersma & Hayes (2001) suggest that the noise starts at 10 for the initial stage, and then revert back to 2 for the consecutive learning stages (p. 80).[34] This is because a learner gets better at filtering out the noise as his proficiency increases, and will not need to perceive every cue perfectly to deduce intended meaning.

The value used at evaluation time is called the selection point and can vary within the category. This is the preferred value of that speaker at that time, i.e. what he perceives. The center point of this range is the value more permanently associated with that constraint, and is what Boersma & Hayes (2001) call the ranking value (p. 47). If two constraint categories overlap due to close values and noise, there are two possible outcomes:

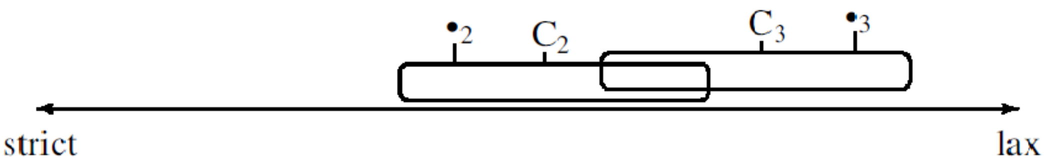

Fig 4.6 Common result: C2 >> C3 (Boersma & Hayes 2001: 48)

In this first and most common outcome, we see that the selection points, 2 and 3, reside in the outer edges of the constraints, thus maintaining the ranking from figure 4.4. The second and more rare outcome is found in figure 4.7 below, where the perception at evaluation time, the real time of which the listener processes the sound signal, results in selection points in the overlapping area.

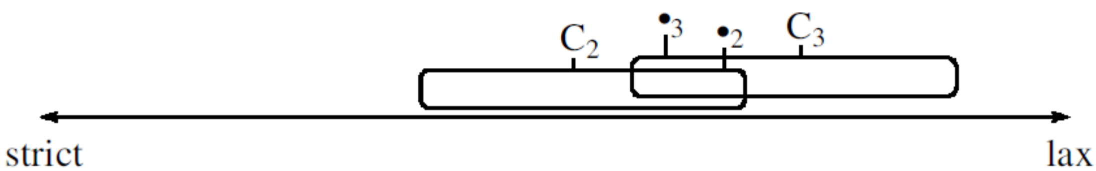

Fig 4.7. Rare result: C3 >> C2 (Boersma & Hayes 2001: 48)

The selection points are now at the other end of the overlapping constraint categories, and selection point 3 outranks selection point 2, providing a different ranking than seen in figure 4.4. This means that we can have more than one output for one input, allowing for free variation.

This method of ranking is highly useful, and necessary, when doing a computer simulation. As there is no computer simulation in the present thesis, the important aspect to bring from

[34] 10 and 2 are values that represent a variable for the chances of misperception. This variable is used in computational analyses.

this presentation into the analyses is that of probability and variation. This will be discussed further in Chapter 5.

4.2.1.2 Escudero's L2LP Model

Escudero (2005, see section 4.1.2 above) proposes a framework for analysis of second language acquisition within Optimality Theory. The full L2LP model was only described for the SUBSET scenario (fewer categories in the L1 than in the L2), and the creation of new categories in a NEW scenario (less categories in the L1 than in the L1) was only described by implementation of new cues.

There are two components to this model; a *perception grammar* containing cue constraints, and a *recognition grammar* with faithfulness and lexical constraints. These interact in the sense that the output of the perception grammar is fed to the recognition grammar as its input, and then the learning step from the recognition grammar alters the ranking of the perception grammar.

The perception grammar takes its input from perceived Hertz values of formants, as presented in section 4.2.1 above. The ranking of the cue constraints is in the present thesis determined by the informant replies. For the value of 2200 Hz, for example, the Mandarin Chinese identified it as a Norwegian /y/ 4 times, and as a Norwegian /ʉ/ 25 times, giving the ranking [2200 Hz] not /y/ >> [2200 Hz] not /ʉ/ in the L2 perception grammar.

[2200 Hz]	[2200 Hz] not /y /	[2200 Hz] not /ʉ/
☞/nʉ/		*
/ny/	*!	

Tableau 4.4 Interlanguage ranking by the Mandarin Chinese informants

The winning candidate, /nʉ/,[35] is what the listener perceives, and this information from the perception grammar is brought to stage two, the recognition grammar. Escudero (2005) has thus created two separate grammars, one with perception and one with lexical recognition.

She also brings with her semantic knowledge to the recognition grammar, marked as such in the input: "Intended: 'new". This is because the Norwegians identified the value of 2200 as

[35] As mentioned in footnote 9, "all the Norwegian vowels discussed in this thesis are long, and will from here on out not be transcribed with the marker of length, [:]."

/ny/,[36] which has the semantic meaning ‘new’. The semantic input represents the knowledge the learner receives from her L2 surroundings and/or the L2 classroom. To be in possession of this knowledge in the input-state is crucial for any learning. In the (incomplete) tableau below, we see that the learner brought with her /nʉ/ from the perception grammar (tableau 4.4 above), together with the semantic knowledge about the intended meaning ‘new’ into the recognition grammar:

/nʉ/ Intended: ‘new’	Constraint	Constraint
Candidate		

Tableau 4.5 An example of the input in the perception grammar after semantic learning

The recognition grammar consists of *lexical constraints* for all candidates, as well as *faithfulness constraints*. The faithfulness constraints, FAITH */vowel1/ → /vowel2/, penalize a change in the vowel quality of vowel1 into the vowel quality of vowel2. If the input from the perception grammar is /ny/, and we have a faithfulness constraint of the type “FAITH */y/ → /ʉ/” a violation will be incurred by the candidate /nʉ/. The ranking of faithfulness constraints relies on perceptual distance between categories, so that the Norwegian ranking can be assumed to be ‘FAITH */ʉ/ →/u/’ >> ‘FAITH */ʉ/ → /y/’. This is because there is more perceptual distance from the mean value of /ʉ/ (1600 Hz) to the beginning of the /u/ category (700 Hz) than to the beginning of the /y/ category (600 Hz), as seen in Chapter 3.

The lexical constraints are naturally ranked below the faithfulness constraints in the beginning, before semantic learning happens, but the faithfulness constraints are demoted at a faster rate than the lexical constraints (*LEX) because “FAITH constraints apply to every perceptual input containing the same vowel while *LEX constraints apply to only one perceptual input.” (Escudero 2005: 227).

A lexical constraint is here of the type ‘*LEX |ny| ‘new’’, meaning that the learner is not to perceive the input as [ny:] because this constraint says that this phonological form does not correspond to the semantic meaning ‘new’. There is one *LEX constraint for every phonology/semantic pairing. In the example below, the listener has heard a Norwegian speaker uttering something that corresponds to the semantic meaning ‘new’ (in L1 Norwegian: /ny/), inserted in the input as “Intended: New”. The listener knows this because of

[36] At 2214 Hz, the native Norwegians identified the stimulus as /ny/ 24 times, and as /nʉ/ only once.

the semantic or real world context. From the perception grammar, however, he brings with him the perception of the input as the phonetic form /nʉ/, which is also inserted in the input.

/nʉ/ Intended: 'new'	*LEX |ny| 'new'	*LEX |nʉ| 'new'	FAITH */ʉ/→ /y/
/ny/ 'new'	*!		*
☞/nʉ/ 'new'		*	

Tableau 4.6 An example of an analysis in the recognition grammar

The 'LEX' constraints are ranked according to perceptual frequency, and this poses a problem for the model, and consequently also for the implementation of the model in the present thesis. Escudero (2005) calculates the ranking based on how many times an L2 token is identified as different categories. For example:

> "*LEX |tʃika| 'girl' [is] ranked higher than *LEX |tʃɪka| 'girl' because tokens of the Spanish word chica are more frequently perceived [by a Dutch native] as /tʃika/ than /tʃɪka/ in a proportion of 79.5% to 20.5%, (p. 222)

For the Norwegian 'new' token , which was identified to be between 2200 and 2700 Hz by the native Norwegians (see Chapter 3), the Mandarin Chinese informants perceived the values as /nʉ/ 50, 6 % of the time, and as /ny/ only 38, 3 % of the time.[37] This means that *LEX |nʉ| 'new' is lowest ranked for the Mandarin Chinese, leading the analysis to confirm the input pair of /nʉ/ = 'new' as in the tableau above. Consequently, there will be no learning step because there is neither a semantic nor phonological error according to the ranking.

Escudero recognizes that such a problem can arise, but hypothesizes that "a continuous promotion and demotion of lexical constraints will occur" until a ranking that allows for the optimal candidate is achieved (p. 234). She adds that this hypothesis could benefit from a validation by the GLA (2005: 235). An attempt was done by Weiand (2007), and she found that Escudero's (2005) method of determining ranking order "failed to yield good learning results" (p. 21). In her conclusion, however, Weiand (2007) found the model satisfactory to some extent and suggests that employing a decreasing learning rate only might lead to a better result.

[37] Also, 10 % for /ni/ and 1, 1 % for /nu/. Percentages are calculated from the overall results of the Mandarin Chinese, found in chapter 3.

It is not this thesis' aim to validate or invalidate the L2LP model, but the recognition grammar seems to need some tweaks before it can be implemented. The recognition grammar is nonetheless assumed as a hypothetical learning step in the discussions and analyses in this thesis, as the current study's informants have gone through varying degrees of learning.

The idea from the L2LP that will be implemented in the analysis in Chapter 5 is that the output of the recognition grammar is fed back into the perception grammar. This creates a perception – recognition – perception cycle. For example, if we assume that a learning step has happened for the value of 2200, i.e. a correct identification by the Mandarin Chinese of Norwegian /ny/ as 'new', it would be fed back to the perception grammar as "Recognition: |ny| 'new'":

[2200 Hz] Recognition: |ny| 'new'	[2200 Hz] not /y /	[2200 Hz] not /ʉ/
/nʉ/		←*
/ny/	*! →	

Tableau 4.7 A hypothetical correct learning step in the Norwegian phonology

Due to the recognition in the input, the learner is aware that this perceived input should be recognized as /y/, and alters the ranking to achieve /y/ as the optimal candidate. The category boundaries for /y/ and /ʉ/ are thus shifted, and the learner is one step closer to establishing native category boundaries.

It is this perception grammar that will be the main component in the analysis in chapter 5, but *a* recognition grammar will also be assumed to have a place in the learning cycle.

4.2.2 A new phonology for the L2: From three categories to four

Before category boundaries of the new L2 categories can be adjusted, as in the model outlined in the previous section, the learner must have some motivation and evidence for adding a new category, here close central rounded, abbreviated /ʉ/.[38] The Mandarin Chinese learner must realize, and implement, a shift in the phonological dimension from a three-category acoustic space in the L1, to a four-category acoustic space in the L2.

[38] The Chinese Mandarin language has, as discussed in Chapter 1, the categories /y/ and /u/ in their L1, so the learning task for these categories is that of adjusting the boundaries, not establish new categories as in the sense of /ʉ/.

A framework for analysis of this shift can be found in Flemming's (2004) Dispersion Theory. According to this theory, it is not the sounds themselves that are perceptually marked, but rather the contrasts between sounds. The general principle is that "contrasts are more marked the less distinct they are." (Flemming 2004: 234). Importantly, Flemming (2004) posits that central vowels are not what is marked, but rather that central vowels are contrasted with front and back vowels. He exemplifies this by comparing the less marked system [i-u] with the more marked systems [i-ɨ] and [ɨ- u], (p. 235) the latter of which are "less than maximally distinct" in terms of F2 (degree of backness), (2004: 236).

The MINDIST (minimal auditory distance) constraints work for maximized auditory contrast. These constraints have the format MINDIST=D:n, where D is the formant (e.g. F2) and n is the *minimal distance* between the contrasting sounds. The candidates are penalized if they do not have a larger contrast than what is given in the MINDIST, and are penalized per contrast. These have a universal, fixed ranking, where MINDIST F1: 1 is always higher ranked than F1: 2, and MINDIST F1: 2 is ranked over F1: 3, etc., so that the smaller the contrast, the more severe the violation, (2004: 239). The distance is derived from the three-dimensional vowel space in figure 4.8 below:

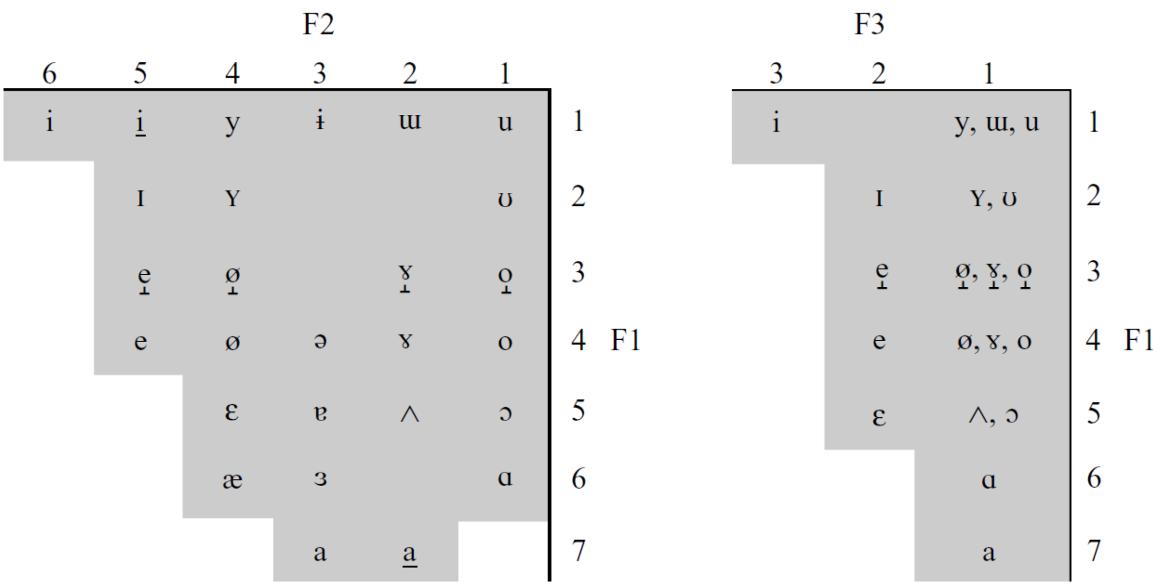

Fig 4.8 A three-dimensional vowel space (Flemming 2004:238).

From Flemming's (2005) analysis on page 242, we derive that the number of distance does not include the vowel that is counted from, e.g. the distance from i to ɨ is 3, violating 'MINDIST F2= 4'. As Flemming (2004) points out, this is a "coarsely quantised" figure (p. 238), but it is found adequate for the present study. For the present study it is important to note that the "central rounded [ʉ] occupies the same position as back unrounded [ɯ]" (2004: 238), which is at place 2 on the F2 scale.

The MINDIST constraints are in conflict with the positive constraint MAXIMISE CONTRASTS, which works for an inventory with a maximum amount of contrastive sounds. This constraint does not penalize candidates, but assigns a checkmark, ✓, for every contrast. The candidates with the least checkmarks are eliminated. Because the MINDIST constraints have fixed ranking, it is thus the ranking of MAXIMISE that decides the difference in the amount of categories in languages.

	MINDIST = F2: 3	MAXIMISE CONTRAST	MINDIST = F2: 4	MINDIST = F2: 5
☞i-u		✓✓		
i-ɯ		✓✓		*!
y-u		✓✓		*!
i-ɨ		✓✓	*!	*
i-ɨ-u	*!	✓✓✓	**	**

Tableau 4.8 A Dispersion Theory ranking that allows for the vowel inventory of /i/ and /u/, (Flemming 2004: 242).

In tableau 4.8 we see that candidate [i-ɨ-u] violates the highest ranked constraint, MINDIST = F2: 3, because [ɨ] and [u] differ by only 2 degrees in the three-dimensional vowel space (see fig. 4.8 above). The MAXIMISE CONTRAST constraint thus has no effect, as the surviving candidates all have two contrasts each.

In the present thesis, the Mandarin Chinese language only allows for the i-y-u distinction, and therefore only asks for three categories in the close acoustic space. Assuming that Mandarin Chinese has the front category /y/, as discussed in Chapter 1, the ranking would be as follows:

	MinDist	Maximise	MinDist	MinDist

	=F2:2	Contrast	=F2:3	=F2:4
y-ʉ-u	*!	✓✓✓	**	***
☞i-y-u		✓✓✓	*	**
ʉ-u	*!	✓✓	*	*
y - ʉ		✓✓!	*	*
i-y-ʉ-u	*!	✓✓✓✓	***	***

Tableau 4.9 The native Mandarin Chinese ranking according to Dispersion Theory.

For Norwegian, however, that allows for 4 categories in the close acoustic space, 'Maximise Contrast' is higher ranked to allow for less distinct contrast in favor of a higher number of contrastive sounds:

	Maximise Contrast	MinDist =F2:2	MinDist =F2:3	MinDist =F2:4
☞i-y-ʉ-u	✓✓✓✓	*	***	***
i-y-u	✓✓✓!	*	*	**
ʉ-u	✓✓!	*	*	*
y - ʉ	✓✓!		*	*

Tableau 4.10 The native Norwegian ranking according to Dispersion Theory.

When the learner becomes aware of the difference in categories, he makes adjustments to his constraint ranking by promoting the constraint that will allow for more contrast at the expense of less distinction:

	MinDist =F2:2	Maximise Contrast	MinDist =F2:3	MinDist =F2:4
y-ʉ-u	*!	✓✓✓	**	***
💣i-y-u		←✓✓✓	*	**
ʉ-u	*!	✓✓	*	*
y - ʉ		✓✓!	*	*
✓i-y-ʉ-u	*!	✓✓✓✓	***	***

Tableau 4.11 The learning step made by the Mandarin Chinese when recognizing the fourth category in the close acoustic space.

This reranking is thus the first learning step that is needed by the Mandarin Chinese while learning Norwegian. It is assumed that the first encounter with the new contrastive category creates a phonological category, here close central rounded,[39] abbreviated as /ʉ/. The consequence of the creation of a new category in the close acoustic space is a reranking in the perception grammar.

4.2.2.1 Implementation of the new category in the perception grammar

In the present thesis, this change in perception grammar is assumed to be the demotion of cue constraints for /ʉ/. This entails that Mandarin Chinese have these cue constraints high ranked in their L1. But do constraints on /ʉ/ exist in Mandarin Chinese, when Mandarin Chinese natives have never experienced or been aware of such a category? The traditional OT view is that all constraints exist in all languages and what separates languages is the ranking of these constraints. The reasoning for this universality is that "since language-particular ranking is in general able to account for languages where a putatively universal constraint does not hold true·, it does not seem necessary to recognize a special class of language-particular constraints." (McCarthy 2002: 11).

Psychologically, however, language-particular constraints and candidate sets can be argued to be motivated. Golston (in Blaho et al., 2007) asks a compelling question: "what serious reason is there to think that my grammar generates and evaluates things I can neither say nor perceive?" (p. 348). Not having the constraints entails that Gen is also restricted, so that it will not generate a candidate that will go uncontested through EVAL due to the absence of the language-particular constraints. This view of language-particular constraints and candidate set is seemingly shared by some researchers of loanword adaption, who do not include the foreign categories to be evaluated in the analyses. One example of this is Kenstowicz & Suchato (2006:935), where /g/ in the input is changed to /k/ in the candidate set because the target language does not have the category /k/. The category change is addressed, and explained by mapping to a native sound that resembles the unknown sound, but this is separate of the OT analysis.

Importantly, OT has never claimed to explain how the brain works, but rather how languages are systematized in universal patterns. As a grammatical model, OT does not explain "the actual processing of linguistic knowledge by the human mind" (Kager 1999: 26). This difference between the grammar and the cognitive performance is, as discussed in this section, not unproblematic.

[39] Ref. the discussion in Chapter 1 about Chinese Mandarin having the close front rounded vowel, /y/, in their L1.

In the present thesis, constraints and candidates are assumed to be universal so that there is no such thing as language-particular constraints or candidate sets. This choice is made for the sake of practical implementation, and the result is that the cue constraints for /ʉ/[40] exist as high ranked constraints in the Mandarin Chinese language. For the Mandarin Chinese, which has close front and close back categories, but not a close central category, constraints banning the perception of close central are necessarily higher ranked than the constraints banning the other two categories. This is because the close central category is unknown and unused, thus marked in the Mandarin Chinese language.

As the category is unknown and unused, one can assume that these cue constraints are simply stored as *ʉ: Do not perceive /ʉ/ in the Mandarin Chinese perception grammar. A high ranking of constraints banning /ʉ/ can also be argued to have its basis in Universal Grammar, where the /ʉ/ is only found in 6 languages, (Maddieson 1984:252). Constraints banning specific categories are found in literature (e.g. *θ in Lombardi 2003:229, but she also notes that this constraint is problematic). However, as presented above, Flemming (2004) argues that it is not the close central category itself that is marked, but the contrast it adds in the perceptual space. As such, the constraint banning the perception of the close central category tells the listener not to perceive a central contrast to the front and back categories, but map all incoming values to either front or back. The cue constraints circumvent this problem by specifying that "so-and-so Hertz value" is not that category, instead of banning the category altogether.

4.3 Summary

The model posited in the present thesis is that the alteration cycle of the perception and recognition grammar in the L2LP can be extended to the MinDist/Max "grammar" in the *initial* L2 phase. This interaction leads to a reranking of the cue constraints for /ʉ/, demoting them below cue constraints for /u/ and/or /y/, as will be further discussed in Chapter 5. After this alteration, only the perception and recognition grammar continue to interact, as the MinDist/Max "grammar" has played its part and is not influential until evidence of more or less categories are presented. The learner will later go on to determine the phonetic distribution of this category in the perception and recognition grammar as posited by Escudero (2005). In the next chapter, Chapter 5, analyses of the present thesis' informant replies will be carried out in accordance with the framework presented above.

[40] The cue constraints can be argued to be of the form "[xxxx Hz] is not /close central rounded/" etc., but are abbreviated to "[xxxx Hz] is not /ʉ /" etc.

Chapter 5 Analysis and discussion of recurring patterns

The analysis and discussion in this chapter is not to be seen as a representative conclusion for the perception of Norwegian vowels by Mandarin Chinese learners. Rather, the collected data is used as a tool to shed light on the theory and identify tendencies in acquisition. Most of the discussion concerning theoretical issues has been addressed in Chapter 4. In this chapter, the recurring patterns and the possible explanations for them within OT and GLA theories will be looked into in more detail. Full analyses are however not possible without following the learners through their learning, as the GLA models need to access both the starting point and the different learning steps to be fully executed.

At the current stage in the learning process, the category boundaries correspond to neither the L1 nor L2 categories for most of the informants. All current GLA theories, to the best of my knowledge, assume semantic learning as the catalyst for second language learning. The results in the present thesis do not necessarily support this (but see discussion in Chapter 6).

Whether or not the learners are exposed to peripheral category tokens is a vital question for subsequent discussion. Will the learners ever be exposed an F2 of 2100 Hz for the semantic meaning "now (old)", pronounced [nʉ], when the production value lies around 1700 Hz (van Dommelen, p.c., see section 1.1.1)? Even when taking into consideration individual variations in pronunciation, the answer is probably no. However, the native Norwegian informants show clear boundaries, especially between /ʉ/ and /y/. The native Norwegians have also been learners of Norwegian once, but still managed to create boundaries which were uniform across different informants (see appendix 6). This indicates that some learning must happen even at peripheral values, and in the present thesis this is assumed to be semantic learning.

As we can see in figure 3.15, repeated as figure 5.1 below, all informants perceive the L2 /ʉ/ in the acoustic space of both the L1 and L2 /y/. Contrarily, /ʉ/ is not perceived in the acoustic space of /u/, neither the L1 nor the L2 category. Furthermore, the /u/ category in Norwegian is quite small, but the Mandarin Chinese informants mostly perceive the L2 /u/ category similar to the significantly larger L1 /u/ category:

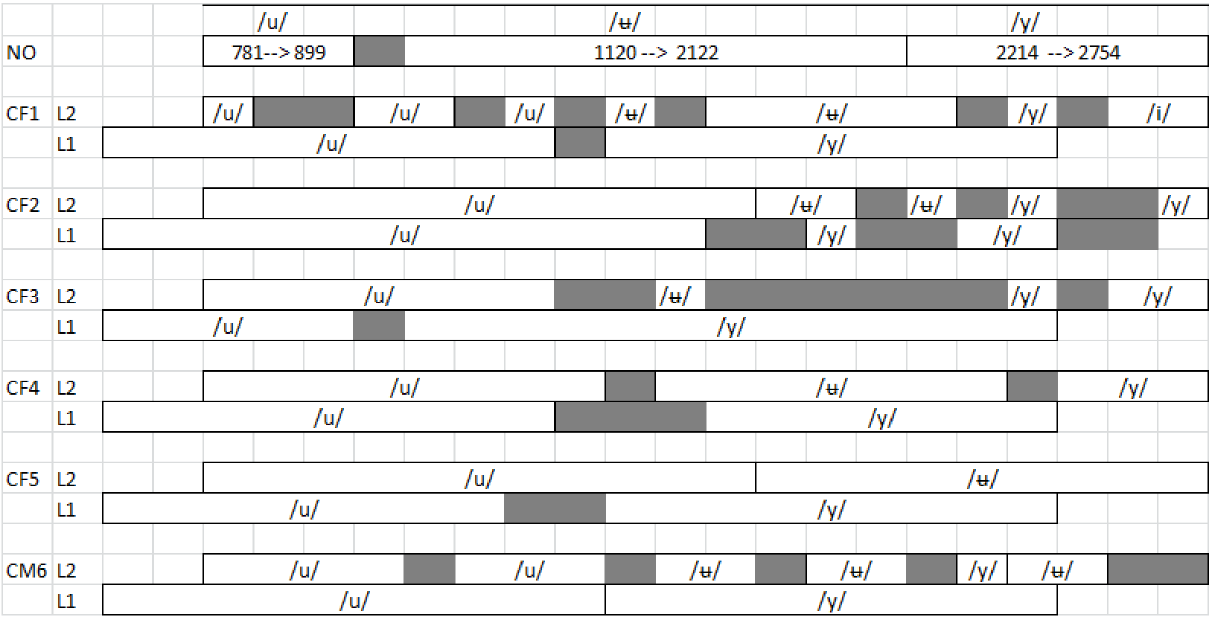

Fig 5.1 A comparison of the overall results of the Native Norwegians (NO) and the Mandarin Chinese native categories (L1) and how they perceived Norwegian categories (L2)

In view of this, there are two primary issues for the GLA theories presented in Chapter 4 in this study: The Norwegian L2 /ʉ/ category is more front for the Mandarin Chinese learners than it is for the Norwegian natives, and the L2 /u/ category is larger than it should be. Semantic learning in a recognition grammar cannot explain this, as will be discussed in this chapter.

Section 5.1 looks at the /ʉ/ - /y/ contrast, and section 5.2 looks at the preservation, and slight expansion of the /u/ category in the L2. The hypothesis posited in the present thesis is that conscious knowledge overrides phonological knowledge in this acoustic space.

5.1 The /ʉ/ - /y/ contrast

Creating a boundary between the Norwegian /ʉ/ and /y/ seems to be the hardest task for the Mandarin Chinese. While most informants have a more or less clear /u/ category in the L2, the acoustic space of /ʉ/-/y/ identifications is riddled with undetermined category boundaries. What learning steps have happened can only be hypothesized without data on the unlearned[41] and the initial interlanguage[42] state of the informants. Only by performing perception tests prior to L2 exposure, and then immediately after being made aware of the new category, can

[41] Unlearned refers to a state where the informants has had no prior exposure to Norwegian.
[42] Interlanguage refers to the initial acquisition state of the L2, where the learner is yet to establish boundaries, but has been made aware of the L2 category inventory.

we know for certain if the /ʉ/ category expands or retracts through learning. Due to the short time frame of a master's thesis, such extensive testing was not possible.

An expansion of the /ʉ/ category would be difficult to explain with the current learning models. It would however be possible to hypothesize that the initial perception of /ʉ/ in most of the acoustic space is a result of overcompensation, i.e. conscious knowledge overriding phonological knowledge. The assumption in the present thesis is for this reason that upon initial exposure to the new category, the Mandarin Chinese chose to perceive /ʉ/ rather than /y/ over a large acoustic space. The informants in the present thesis differ as to how large this acoustic space is. It is reasonable to assume that the peripheral values of /y/,e.g. from 2500 Hz, would not be perceived as /ʉ/, even initially. This is discussed further in sections 5.1.1 and 5.1.2 below.

5.1.1 Perception of the /ʉ/ category as front

In the present thesis, the learners have created the new close central rounded category in the L2 phonology. They know they *should* perceive this category, abbreviated /ʉ/, in the L2, and are assumed to react by *choosing* to perceive this /ʉ/ to a high degree (see figure 3.15, repeated as fig 5.1 above). What was suggested in Chapter 4, based on the results in Chapter 3, is that the conscious knowledge overrides phonological knowledge when they are taught a new category. By experiencing that there is a contrastive category in the L2 inventory that is not in the L1 inventory, eager learners will do their best to perceive this contrast at the expense of their native filter (see Chapter 4).

The phonological knowledge should, according to theories presented in Chapter 4, result in an initial mapping of /ʉ/ to L1 categories (see section 5.1 above). Through semantic learning of each formant value in the recognition grammar, as outlined in section 4.2, the learners would be taught that the value they perceived as e.g. /y/ in their L1, should instead be perceived as /ʉ/ in the L2. The input of 1900 Hz in the tableau below is 200 Hz from the average production value of Norwegian /ʉ/ at 1700 Hz:

1900 Hz Recognition: /ʉ/	[1900 Hz] is not /ʉ/	[1900 Hz] is not /u/	[1900 Hz] is not /y/
💣ny			←*
✓nʉ	*! →		
nu		*!	

Tableau 5.1 A hypothetical learning step in the Norwegian L2 perception grammar.

Consequently, one would expect that the new category, here /ʉ/, is acquired gradually, as there is no close, central category in the L1. Contrarily, the close front categories, labeled /y/ by UPSID, are quite similar in the two languages (see chapter 1). The acquisition task of the L2 /y/ category should then only be one of adjustment of boundaries (as per discussion about peripheral values above). Importantly, it is evidence from the L2 environment about wrong perception that spawns a learning step in both the recognition and perception grammar (see section 4.2). The learners would not receive evidence that would enforce a demotion of the cue constraints for /ʉ/ below the cue constraints for /y/ from the values of 2200 Hz and upwards in Norwegian, as these values are considered a Norwegian /y/. No Norwegian input would invoke a semantic learning step that asked the Mandarin Chinese to perceive /ʉ/ instead of /y/, i.e. 'now (old)' instead of 'new' for those values.

2400 Hz Recognition: /ʉ/	[2400 Hz] is not /ʉ/	[2400 Hz] is not /u/	[2400 Hz] is not /y/
💣ny			←*
✓nʉ	*! →		
nu		*!	

Tableau 5.2 A theoretically impossible learning step in the Norwegian L2 perception grammar.

The tableau above illustrates a hypothetical learning step that would allow for semantic learning to trigger the incorrect perception of 2400 Hz as /ʉ/ in the L2. As discussed above, this learning step is theoretically impossible in Norwegian.

Still, the Mandarin Chinese learners perceive the new category /ʉ/ often in the acoustic space of L2 /y/. This cannot be explained by semantic learning, and it cannot be explained by the L1 phonology because there is no /ʉ/ in the L1. The reason for this seemingly unmotivated perception of /ʉ/ is thus hypothesized to be because conscious knowledge overrides phonological knowledge. An illustration of this phenomenon is illustrated by an analysis of informant CF4 in the next section.

5.1.1.1 Informant CF4

CF4 is one of the informants who identified /ʉ/ in the acoustic space of both the L1 and L2 /y/ category. She is not far off from establishing the correct boundary between /y/ and /ʉ/, but she still perceives /ʉ/ further up the continuum than she should, ref. figures 3.2 and 3.12, repeated as figures 5.2 and 5.3 below:

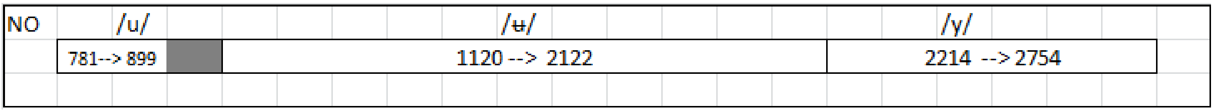

Fig 5.2 The native Norwegian category boundaries

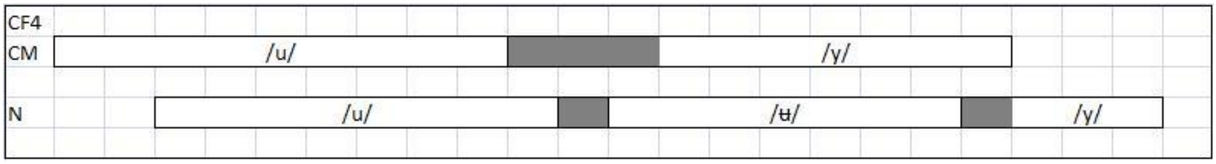

Fig 5.3 A visual figure of CF4's category boundaries in Mandarin Chinese (CM) and Norwegian (N)

Figure 5.3 shows that CF4 has managed to establish three categories in Norwegian, /u/, /ʉ/ and /y/. There is one confused area between each of the L2 boundaries, something which is natural for overlapping areas. However, her category boundaries in the L2 are not correct. In this section, the category boundary between /ʉ/ and /y/ will be analyzed.

CF4's clear category boundary between /ʉ/ and /y/ is located at 2500 Hz (see table 5.2 below), while the native Norwegian boundary is at 2200 Hz (see fig 5.2 above). Her L2 boundary does not resemble that of the boundary between her L1 /u/ and /y/, which is at 1800 Hz (see table 3.10 & 3.11, reproduced as tables 5.1 & 5.2 below).

Mandarin Chinese		
F2	/u/	/y/
631	5	
758	5	
814	5	
984	5	
1033	5	
1188	4	1
1262	4	1
1329	4	1
1412	5	
1497	2	3
1600	2	3
1703	2	3
1861		5
1936		5
2074		5
2172		5
2259		5
2385		5
2547		5

Norwegian				
F2	/u/	/ʉ/	/y/	/i/
781	5			
899	5			
1012	5			
1120	5			
1237	5			
1324	5			
1402	5			
1518	5			
1652	3	2		
1793	1	4		
1899		5		
1956	1	4		
2053		5		
2122		5		
2214		5		
2335		4	1	
2433		3	2	
2518		1	4	
2650			5	
2754			5	

Tables 5.1 & 5.2 The native responses to the Mandarin Chinese perception test (left), and the Mandarin Chinese responses to the Norwegian perception test (right) by informant CF4.

If we assume that most L1 /y/ values were mapped to L2 /ʉ/ in the initial stage, CF4 has successfully reduced her /ʉ/ category in the upper end of the continuum. Her next learning step is to no longer perceive 2300 Hz as /ʉ/.[43,44] For this value, then, she has gone from an L1 ranking where the cue constraint for /ʉ/ is highest ranked due to markedness (see section 4.2.2.1), and the cue constraint for /y/ is lowest ranked because her responses to this stimuli in the L1 test favored /y/ over /u/ (see table 5.1 above). This gives the following analysis of 2300 Hz in her L1 perception grammar:

[43] The native Norwegians identified this value as /y/ in 24/25 replies.

[44] 2400 Hz is not determined; see section 5.3 below for analyses on this phenomenon.

2300 Hz	[2300 Hz] is not /ʉ/	[2300 Hz] is not /u/	[2300 Hz] is not /y/
☞ny			*
nu		*!	
nʉ	*!		

Tableau 5.3 The L1 ranking for the value of 2300 Hz for CF4

The optimal candidate in CF4's L1 perception grammar is thus /ny/ for the value of 2300 Hz. According to second language acquisition theories, this is the filter through which she should perceive the L2 sounds.

At the time of the perception test, when CF4 had lived in Norway for 2 years and attended level 1 of the Norwegian course, her L2 ranking for the value of 2300 Hz was: [2300 Hz] is not /u/ >> [2300 Hz] is not /y/ >> [2300 Hz] is not /ʉ/. The cue constraint for /u/ is highest ranked because she had no /u/ responses to this stimulus, as seen in table 5.2 above. The cue constraint for /ʉ/ is lowest ranked because she favored this to /u/ four to one (see discussion on ranking of cue constraint in section 4.2.1.2)

2300 Hz	[2300 Hz] is not /u/	[2300 Hz] is not /y/	[2300 Hz] is not /ʉ/
ny		*!	
☞nʉ			*
nu	*!		

Tableau 5.4 The Norwegian L2 ranking for the value of 2300 Hz for informant CF4.

The optimal candidate in CF4's L2 perception grammar is /nʉ/ for the value of 2300 Hz. How did this happen? As discussed above, there would be no learning step where she would be corrected from /y/ perception to a /ʉ/ perception for this value in the L2, because a value of 2300 Hz in Norwegian is considered a /y/.

In tableau 5.5 below, "Recognition" denotes the learning she brought with her from the recognition grammar (see section 4.2). According to the L2LP, semantic recognition in the recognition grammar is the only learning step that could explain why she would rerank the constraints in the perception grammar to allow for a /ʉ/ perception in the L2.

2300 Hz Recognition: /ʉ/	[2300 Hz] is not /ʉ/	[2300 Hz] is not /u/	[2300 Hz] is not /y/
💣ny			←*
✓nʉ	*! →		
nu		*!	

Tableau 5.5 An incorrect learning step in the Norwegian perception grammar.

The above analysis is incorrect as the results in the present thesis (Chapter 3) show that 24/25 native Norwegian responses were in favor of /y/, thus this input should have the output /ny/ in the Norwegian perception grammar. Based on this, it is highly unlikely that she would have been corrected from a /y/ perception to a /ʉ/ perception by a Norwegian native or by other evidence.

The other possibility is that CF4 demoted the cue constraints for /ʉ/ as soon as she realized there was such a category in the L2 (see discussion in section 4.2.2.1). This realization, as discussed in section 4.2.2, is formalized in Flemming's (2004) Dispersion Theory. The learning step is shown in tableau 4.11, reproduced as tableau 5.6 below:

	MinDist =F2:2	Maximise Contrast	MinDist =F2:3	MinDist =F2:4
y-ʉ-u	*!	✓✓✓	**	***
💣i-y-u		←✓✓✓	*	**
ʉ-u	*!	✓✓	*	*
y - ʉ		✓✓!	*	*
✓i-y-ʉ-u	*!	✓✓✓✓	***	***

Tableau 5.6 The learning step made by the Mandarin Chinese when recognizing the fourth category in the close acoustic space.

The hypothesis for CF4, and the other informants, is that when she initially realizes that there is a category /ʉ/ in the L2, she contrasts this with the L2 /y/ (see section 5.2 below for discussions on the /ʉ/ - /u/ contrast). She is however not able to contrast between them properly. The information she receives from perception/the classroom is that in the space where she would perceive /y/ in her native language, there is both /y/ and /ʉ/ in the L2. The aforementioned hypothesis about conscious knowledge overriding phonological knowledge triggers a constraint reranking that allows for the new, central, category to be perceived to a higher degree than the known, front category. Most "[xxxx Hz] is not /ʉ/" are thus demoted below most “[xxxx Hz] is not /y/”. As a result, she rejects her L1 phonology, the filter that she is supposed to perceive the new phonology through (as discussed in section 4.1), and replaces it with the conscious knowledge that there should be something else.

How exactly this replacement is translated to an OT analysis, is as of now unclear, but an additional input of the kind “Perceive: /ʉ/”[45] would show the presence of the Dispersion Theory in the perception grammar:

2300 Hz Perceive: /ʉ/	[2300 Hz] is not /ʉ/	[2300 Hz] is not /u/	[2300 Hz] is not /y/
💣ny			←*
✓nʉ	*! →		
nu		*!	

Tableau 5.7 A hypothetical learning step when adding the extra information of “Perceive” in the input.

The interaction between Dispersion Theory and perception grammar is in the present thesis hypothesized to happen only once, in the initial phase where the new category is *first* discovered. This input would then apply to all the stimuli the informant would identify as /ʉ/. After this first appearance in the perception grammar, “Perceive: /ʉ/” would become replaced by the “Recognition” input from the recognition grammar. However, this suggestion is not tested nor found theoretically sound.

Another possible explanation for why the results show that not all cue constraints for /ʉ/ are demoted below the cue constraints for /y/ is flaws with the perception test. The uppermost values, 2500 to 2700 Hz, of the L2 continuum are identified as /y/ by CF4. Only one of these values, 2500 Hz, is included in the Mandarin Chinese perception test, and they did not have the option “ni” in their L1 test either (see Chapter 2 for discussion on this issue). Consequently, these values are difficult to analyze. It is worth noting that the Mandarin Chinese production value for /i/, as given by Zee and Lee (2001), is at 3000 Hz. It might be that the perceptual category of /i/ spans down to 2500 Hz, but without data on this, it is theoretically equally possible that it does span that far as it is that it does not.

[45] Similar to the “Recognition” input from the recognition grammar, or the “Intended” input in the recognition grammar.

Due to orthography (see Chapter 1), it is also possible that the informants did not see the “ny” option in the L1 test as equal to the “ny” option of the L2 test, i.e. they did not identify them as belonging to the same category. If the informants did see the L2 <nu> option, pronounced [ʉ], as equal with the L1 <ny> option, pronounced [y], it would explain why there is such a large identification of /ʉ/ in the front acoustic space. On the other hand, if the front category was represented orthographically as <nü> in the L1 test, the probability would be greater for the possibility of the L1 and L2 front category being perceived as two different categories by the informants. Conclusively, it is of my opinion that until further testing unveils the answers to the above questions, the hypothesis of conscious knowledge overriding phonological knowledge is a viable explanation.

5.1.2 /ʉ/ replaces /y/

The results of informant CF5 show that she has not been able to distinguish between the L2 /ʉ/ and /y/, as seen in figure 3.13, reproduced as figure 5.4 below:

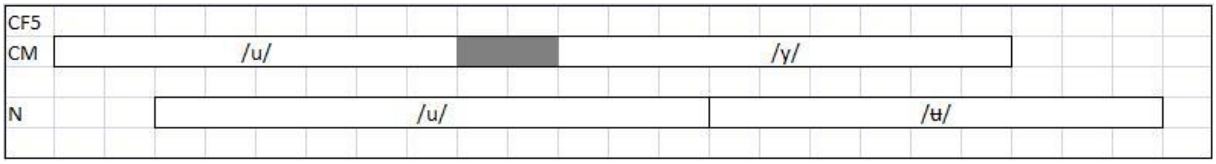

Fig 5.4 A visual figure of CF5’s category boundaries in Mandarin Chinese (CM) and Norwegian (N)

Instead, she has created what can be analyzed as a composite category (see section 4.1.2), where she assimilates the two categories, L2 /y/ and /ʉ/, to, surprisingly, the unknown category, /ʉ/. If we consider the hypothesis of conscious knowledge overriding phonological knowledge, this is not as surprising: She knows she should perceive this /y/-/ʉ/ contrast, but she is unable to. She thus chooses the unknown category, because she knows it is supposed to be in the upper end of the continuum, she just does not know exactly where.

As seen in table 3.12, reproduced as table 5.4 below, she only perceives /ʉ/ in the upper end of the continuum in the L2, and does not have a single identification of a front category, neither /y/ nor /i/. Contrarily, in the L1 acoustic space, she identifies a front category, as seen in table 3.11, reproduced as table 5.3 below:

Mandarin Chinese		
F2	/u/	/y/
631	5	
758	5	
814	5	
984	5	
1033	4	1
1188	5	
1262	4	1
1329	4	1
1412	3	2
1497	2	3
1600		5
1703		5
1861		5
1936		5
2074		5
2172		5
2259		5
2385		5
2547		5

Norwegian				
F2	/u/	/ʉ/	/y/	/i/
781	5			
899	5			
1012	5			
1120	5			
1237	5			
1324	5			
1402	5			
1518	5			
1652	5			
1793	4	1		
1899	5			
1956	1	4		
2053		5		
2122	1	4		
2214		5		
2335		5		
2433		5		
2518		5		
2650		5		
2754		5		

Tables 5.3 & 5.4 The native responses to the Mandarin Chinese perception test (left), and the Mandarin Chinese responses to the Norwegian perception test (right) by informants CF5.

This informant is a level 1 student, but has lived in Norway for 4 years. Thus, there is a possibility that she has had sufficient Norwegian input in her daily life to equal that of formal training, but there is also the possibility that she shows the initial stage of acquisition because she has had little formal training. The latter possibility would indicate that there is an initial /ʉ/ identification of the upper end of the continuum. It is nonetheless an interesting result which shows that CF5 does not consider her L1 /y/ equal to the L2 /y/. Unfortunately, this result also raises questions about the perception test itself, and we cannot draw a decisive conclusion from these results.

5.1.3 The contrast between /ʉ/ and /y/ not established

Many of the informants have undetermined areas throughout their L2 continuum. One of them, informant CF3, has a considerable amount of these in the upper end of the L2 continuum. As seen in table 3.9, reproduced as table 5.6 below, CF3's response to the /ʉ/ - /y/ contrast is at chance level from 1900 to 2300 Hz:

Mandarin Chinese		
F2	/u/	/y/
631	5	
758	5	
814	5	
984	5	
1033	5	
1188	3	2
1262	1	4
1329		5
1412		5
1497		5
1600		5
1703		5
1861		5
1936		5
2074		5
2172		5
2259		5
2385		5
2547		5

Norwegian				
F2	/u/	/ʉ/	/y/	/i/
781	5			
899	5			
1012	5			
1120	5			
1237	4	1		
1324	4	1		
1402	4	1		
1518	2	3		
1652	3	2		
1793	1	4		
1899		2	3	
1956		2	3	
2053		3	2	
2122		2	3	
2214		3	2	
2335		2	3	
2433		1	4	
2518	1	1	3	
2650			5	
2754			5	

Tables 5.5 & 5.6 The native responses to the Mandarin Chinese perception test (left), and the Mandarin Chinese responses to the Norwegian perception test (right) by informants CF3.

CF3 alternates between preferring /ʉ/ or /y/ to these stimuli,[46] and has not established a clear boundary until 2600 Hz. In a traditional OT analysis, not establishing a boundary would mean that the cue constraints for these values are not internally ranked, i.e. [1900 Hz] is not /ʉ/,

[46] The /u/ - /ʉ/ contrast will not be addressed here.

[1900 Hz] is not /y/. The cue constraint for /u/, which is not perceived for the value of 1900 Hz, would be highest ranked. This analysis would leave us with two optimal candidates:

1900 Hz	[1900 Hz] is not /u/	[1900 Hz] is not /ʉ/	[1900 Hz] is not /y/
☞/ʉ/		*	
/u/	*!		*
☞/y/			

Tableau 5.8 The ranking of 1900 Hz in the L2 by informant CF3

If we were to do a computational analysis and employ the SOT method as presented in section 4.2.1, the constraint '[1900 Hz] is not /ʉ/' would hold a higher value than the constraint '[1900 Hz] is not /y/' because /y/ was preferred over /ʉ/ 3 to 2. As such, they are not equally ranked, but close enough to overlap if the noise value is sufficient. This leads to variation in the perception, where it is the selection points and not the ranking value that is decisive for perception. As we are not doing a computational analysis, we will not delve further into this aspect of the analysis, but conclude that CF3 is largely unable to separate /y/ from /ʉ/.

5.2 The /u/ - /ʉ/ contrast

Common to all Mandarin Chinese informants is the large /u/ category in the L2, and for many of them, the L2 /u/ category mirrors that of the L1. This indicates that the L2 /u/ category has been assimilated, i.e. equaled, to the L1 /u/ category (see section 4.1). As this large /u/ category is shared by all informants, the analysis will be based on the overall results, as seen in figure 3.8, repeated as figure 5.5 below:

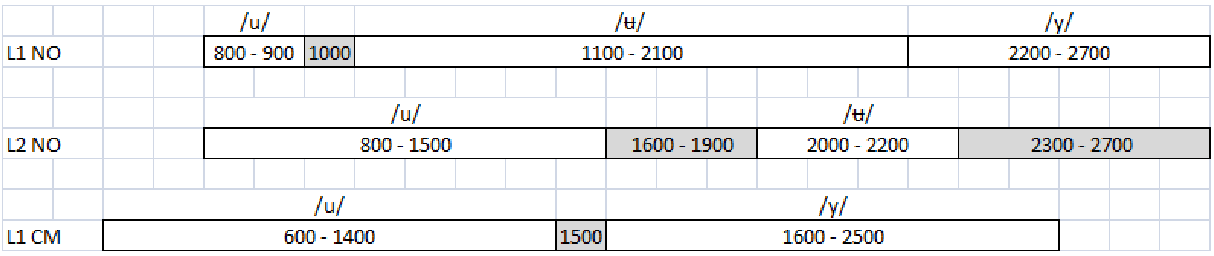

Fig 5.5 The overall results of the native Norwegian category boundaries (L1 NO, top), the Mandarin Chinese perception of the Norwegian category boundaries (L2 NO, middle), and the Mandarin Chinese category boundaries (L1 CM, bottom)

As seen in figure 5.5 above, the native Norwegian /u/ category is quite small. One hypothesis could be that 'Faith */u/ →/ʉ/' (see section 4.2) is ranked very high in the recognition grammar, thus blocking the semantic learning of /ʉ/. If we assume values as according to SOT (section 4.2), we can assume that 'Faith */u/ →/ʉ/' has a very high value so that its demotion from semantic evidence is slow.

An explanation for this high ranking might be found in the teaching methods. One of the teachers at the Norwegian course for Foreigners at NTNU, Kjell Heggvold Ullestad (p.c.), says that he teaches students the /ʉ/ category by contrasting it with the /u/ category by the front/back dimension. The students are told to pronounce a /u/, hold it, and then press the tongue forward until they reach the correct position for the pronunciation of the /ʉ/. It is possible to imagine that this method may lead to the students' awareness of /u/ definitely not being an /ʉ/, thus preserving what they identify as an /u/ in their L1. This explanation would also be in accordance with conscious knowledge overriding phonological knowledge, as the students are instructed to contrast /ʉ/ with /u/. However, the high ranking of FAITH has a phonological basis, and is undoubtedly tied to perception of the L2 through the L1 filter.

5.3 A comparison between beginner students of Norwegian and advanced students of Norwegian

The informants' different time of residence in Norway prevent creating uniform "beginner" and "advanced" groups for comparison, but the level 1 students are here considered "beginners" and the level 3 and 4 students are considered "advanced". This grouping results in small groups of only 3 and 2, which allows for individual identification patterns to emerge to a higher degree than when combining the entire informant base. The discussion below is therefore not in-depth, and is not considered to be representative of the differences between the levels because it holds little to no statistical value.

As seen in figure 5.6 below, the three "beginner" students have some identification of /ʉ/ through most of the continuum, though most of the lower value identifications are attributed to informant CF1. The only clear category for /ʉ/ is between 1900 and 2200 Hz. The /u/ category is clear up until 1500 Hz, and perceived up until 2100 Hz, though barely. The /y/ category is never the dominant choice, but there are identifications from 1900 to 2700 Hz. The beginner students also identify an /i/ category at the upper values, though not as a clear category.

Norwegian				
F2	/u/	/ʉ/	/y/	/i/
781	15			
899	11	4		
1012	12	3		
1120	13	2		
1237	12	3		
1324	13	2		
1402	13	2		
1518	13	2		
1652	9	6		
1793	7	8		
1899	7	8		
1956	1	12	2	
2053		13	2	
2122	1	13	1	
2214		13	2	
2335		8	7	
2433		10	5	
2518		10	4	1
2650		6	4	5
2754		7	3	5

Fig 5.6 The level 1 Chinese informants' replies to the Norwegian perception test

The "advanced" group only consists of two informants. [47] They have, as seen in figure 5.7 below, identified /ʉ/ from1600 to 2600 Hz, though only 1900 to 2200 is a clear /ʉ/ category. /u/ is identified as far up as 2200 Hz, though only just from 1900 Hz. The advanced students have a very clear /u/ category up until 1500 Hz, with no other identifications in that area. The area from 2300 Hz is largely undecided, but the /y/ category is the most identified category. There are also a few /i/ identifications from 2300 Hz, but no clear category or majority identifications for any stimuli.

[47] Informant CF3 is a level 2 student, thus belonging to neither the "beginner" nor "advanced" group. Her results are therefore not included. This makes for uneven groups for comparison.

Norwegian				
F2	/u/	/ʉ/	/y/	/i/
781	10			
899	10			
1012	10			
1120	10			
1237	10			
1324	10			
1402	10			
1518	10			
1652	7	3		
1793	5	5		
1899	4	6		
1956	2	8		
2053	1	9		
2122	2	7	1	
2214	1	9		
2335		6	3	1
2433		3	6	1
2518		1	7	2
2650		1	7	2
2754			9	1

Fig 5.7 The level 3 and 4 Chinese informants' replies to the Norwegian perception test

These results indicate that /ʉ/ is reduced at both ends of the continuum through learning, and that /y/ becomes a stronger category as learning progresses. However, as pointed out above, these results are based on very limited data and hold little statistical or analytical value.

5.4 Discussion and summary

The hypothesis posed in the present thesis is that conscious knowledge overrides phonological knowledge by perceiving /ʉ/ to a higher degree than what is found phonological evidence for in either the L1 or L2 phonology. To revise this hypothesis into empirical evidence, more stages of the Mandarin Chinese learning process would have to be documented. A test of the initial, unlearned,[48] mapping of the Norwegian sounds to the Mandarin Chinese categories, and then the initial, learned,[49] identification of /ʉ/ in the L2 phonology would shed further light on this issue. By knowing these results, we could see if the Mandarin Chinese initially replaced the entire /y/ category with /ʉ/ in the L2, and then proceeded to adjust the category

[48] Unlearned in this sense refers to no prior exposure of Norwegian.
[49] Learned in this sense refers to exposure to Norwegian.

boundaries to make room for an L2 /y/. Alternatively, /ʉ/ has spread through learning, from being placed acoustically in a small space to being identified in the majority of the L2 close front acoustic space. This latter alternative would have consequences for the L2LP model because such a spreading would entail incorrect semantic learning.

The same scenario applies to the /u/ category: If the Mandarin Chinese initially, unlearned, mapped some of the values for the L2 /ʉ/ to /u/, followed by initial, learned, placement of /ʉ/ somewhere in the central area, the results seen above would entail unmotivated spreading of the /u/ category. If, on the other hand, there were no initial mapping of /ʉ/ to L1 /u/, and the /ʉ/ was initially perceptually placed in the upper area of the continuum, assimilation would be the reason for the large /u/ category.

Looking at dialectal differences, where it was established in section 3.1.1 that no Norwegian informant had a /u/ category above 1300 Hz, and no /y/ category below 2200 Hz, they seemingly make no difference for the results. That is, dialectal differences are not the causes for the erroneous category boundaries perceived by the Mandarin Chinese. The overall results of the Mandarin Chinese show that they perceive /u/ much further up the continuum than 1300 Hz, with a clear /u/ category up until 1500 Hz, and perception of /u/ as far up as 2200 Hz. The Norwegian /y/ category is, as already discussed, not established. It is, however, perceived from 1900 Hz. None of the Norwegian informants perceived /y/ at this value.

Lastly, when taking a look at the Norwegian production values of the categories, as presented in section 1.1.1, we see that the Mandarin Chinese do fairly well for /u/ and /ʉ/. Many of the Mandarin Chinese informants almost correctly identified the “prototype” production value of the Norwegian /ʉ/ at 1707 Hz (van Dommelen, p.c.). Three of them perceived /ʉ/ at 1793 Hz, and one informant at 1652 Hz (see table 3.3). The prototype production value of /u/ is 781, and all Mandarin Chinese informants identified the stimulus 781 Hz correctly as /u/. Contrarily, under half of the replies to the stimulus closest to the prototype value of /y/ at 2367 Hz (van Dommelen, p.c.) are correct. At stimulus 2335 Hz, there were 16 identifications of /ʉ/, 1 identification of /i/ and 13 identifications of /y/. Based on identification values, then, it is clear that the Mandarin Chinese have problems identifying /y/ correctly, and not /ʉ/.

Chapter 6 Summary and conclusion

The results in the present thesis indicate that the Mandarin Chinese students overcompensate in the perception of the newly acquired category, /ʉ/. This is hypothesized to be a result of conscious knowledge, i.e. the knowledge that "there should be a category /ʉ/ somewhere in this acoustic space", overriding phonological knowledge.[50] This applies to both L1 and L2 phonological knowledge, where the new L2 category, /ʉ/, is perceived in the acoustic space of both the L1 and the L2 /y/ category. Consequently, neither L1 filtering nor L2 semantic learning can explain the identifications of /ʉ/ further up the continuum than 2100 Hz.

One of the questions posed in Chapter 1 was to what degree the categories differed perceptually between Norwegian and Mandarin Chinese. As seen in Chapter 3, the differences between the L1 /u/ and the L2 /u/ perceptual categories are substantial. This has led to the perception of L2 /u/ in the acoustic space of the Norwegian /ʉ/. The explanation for this can be assimilation of the L2 /u/ to the L1 /u/, or that the magnet effect of the /u/ category affects the boundary adjustment between /u/ and /ʉ/. Alternatively, the teaching method of contrasting /ʉ/ with /u/ can result in conscious knowledge, i.e. do not perceive /ʉ/ where there is /u/, overriding the phonological knowledge given to the learner through semantic learning. Which of the two options apply for the Mandarin Chinese learners must be investigated further with a perception test prior to learning Norwegian. Due to the limited time frame of a master's thesis, such a test was not carried out in the present study.

The results presented in this thesis must be viewed in light of the method used and the problems that arose along the way. This study has shown that applying the correct method is crucial for extracting viable results for analysis. Though every choice made was discussed, debated and made on theoretically sound reasons, the extracted results show that another method is preferable. A pilot study where different methods are applied should be carried out prior to the main study, to ensure that the chosen method suits the informants of the language that is being tested. This is not to say that the results in the present thesis are wrong, but there is a possibility that orthography, as well as the differences in the amount of stimuli and options, in the perception tests can be sources of error in the results.

---

[50] An article by Eckman et al. (2013) on hypercorrection was published in the July 2013 edition of *Second Language Research.* Unfortunately, this article was discovered shortly before the submission of this thesis, thus the discussions in the present thesis did not benefit from Eckman et al.'s insights. The findings on hypercorrection in the acquisitions of L2 phonemic contrast seemingly conclude with similar remarks as what is shown in the present thesis.

Future studies in perceptual acquisition of L2 category boundaries should attempt to examine more stages of learning. Most important are the first stages in the learning process, both prior to exposure and after first exposure. That way, the researcher is able to determine in which direction the categories expand or retract. Furthermore, for experiments where the two languages use different orthographic systems, special care must be taken to ensure that the informants are given the ability to give correct answers. One suggestion is to utilize pictures instead of orthography in the perception tests to ensure that there is no confusion as to what sound corresponds to which alternative on the computer screen. The choice of test words that carry meaning is in that case crucial.

From a teaching perspective, the discovery that Mandarin Chinese students experience difficulties distinguishing between /ʉ/ and /y/ comes as no surprise. What these results may contribute to teaching, however, is that /ʉ/ is misperceived as /u/ quite far up the continuum, and not only at the peripheral values. The Mandarin Chinese informants have a significant identification[51] of /u/ as far up as 1900 Hz in the L2, while the last perception of /u/ by any Norwegian informant was a mere two identifications at 1500 Hz. For teaching assets such as CALST, information about common misperceptions could be included. For example, the Mandarin Chinese L1-L2 map could include the information that the native /u/ is not the same as the Norwegian /u/, though it seems like it on the map. It is a possibility that providing a visual figure of the differences between the languages' perceptual categories can aid learners in establishing the new category correctly in their acoustic space. Figure 6.1 below is an example of how this can be done, though it should be tweaked for such a purpose:

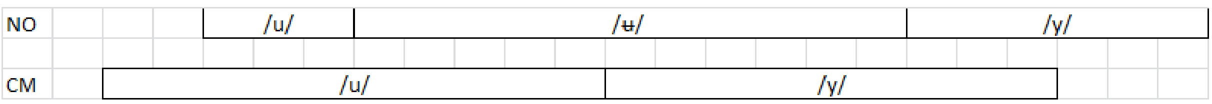

Fig. 6.1 An example of how perceptual ranges could be presented in an L1-L2 map

For more extreme cases, like German as discussed in section 1.1, an L1-L2 map that shows the language-specific placement of the categories' production values is a good starting point.

[51] 11 identifications, see table 3.3.

7. References

Albersten, E. (2008). *Persepsjon av norsk som andrespråk: Oppfattelsen av vokallengde-kontrast mellom norske, fremre, trange og halv-trange vokallyder hos morsmålsbrukere av fransk, spansk og tysk.* Master's thesis, NTNU, Trondheim.

Ashby, M., & Maidment, J. (2005). *Introducing phonetic science*. Cambridge University Press.

Best, C. T., Hallé, P. A., Bohn, O. S., & Faber, A. (2003). Cross-language perception of nonnative vowels: Phonological and phonetic effects of listeners' native languages. In *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 2889-2892).

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. *The Journal of the Acoustical Society of America,109*, 775.

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. *The development of speech perception: The transition from speech sounds to spoken words, 167,* 224.

Blaho, S., Bye, P., & Krämer, M. (Eds.). (2007). *Freedom of analysis?* (Vol. 95). Walter de Gruyter.

Boersma, P., & Weenink, D. (2012). Praat: doing phonetics by computer [Computer program], Version 5.3. 32., retrieved August 2012 from http://www.praat.org/

Boersma, P. (2004). A stochastic OT account of paralinguistic tasks such as grammaticality and prototypicality judgments. *Unpublished manuscript, University of Amsterdam.* Retrieved October 2012 from http://uvafon.hum.uva.nl/paul/papers/ParalinguisticTasks.pdf.

Boersma, P., & Hayes, B. (2001). Empirical tests of the gradual learning algorithm. *Linguistic inquiry*, *32*(1), 45-86.

Boersma, P. (1999). Optimality-theoretic learning in the praat program. In *Proceedings of the Institute of Phonetic Sciences of the Univ. of Amsterdam*(Vol. 23, pp. 17-35).

Boersma, P. (1997). How we learn variation, optionality, and probability. In *Proceedings of the Institute of Phonetic Sciences of the University of Amsterdam* (Vol. 21, pp. 43-58).

Boersma, P., Escudero, P., & Hayes, R. (2003, August). Learning abstract phonological from auditory phonetic categories: An integrated model for the acquisition of language-specific sound categories. In *Proceedings of the 15th international congress of phonetic sciences* (Vol. 1013, p. 1016). Barcelona.

CALST (2012): ”L1-L2 Mapping», http://calst.hf.ntnu.no/L1-L2map, NTNU. Trondheim, retrieved October 2012.

Database for statistikk om høgre utdanning (2012):

http://dbh.nsd.uib.no/dbhvev/student/utenlandske_rapport.cfm?vkode=x&brukersort=to&viskode=0&nullvalue=-&landkode=x&studkode=x&progkode=x&semester=1&sti=landkode,studkode,progkode&insttype=x&arstall=2012&instkode=1150&finans=total&fakkode=x&ufakkode=x&beregning=Totalt.antall&valgt_sti=Norges%20teknisk-naturvitenskapelige%20universitet&grupperingstring=a.arstall&sti_hele=instkode,landkode,studkode,progkode&sti_valgt=instkode,landkode,studkode,progkode, retrieved September 2012.

Duanmu, S. (2007): *Phonology of Standard Chinese*, Oxford University Press, Oxford.

Eckman, F. R, Iverson, G. K, & Song, J. Y. (2013). The role of hypercorrection in the acquisition of L2 phonemic contrasts. *Second Language Research*, *29*(3), 257-283.

Escudero, P. (2005). *Linguistic perception and second language acquisition: explaining the attainment of optimal phonological categorization*. Doctoral dissertation, Utrecht University.

Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. *Laboratory phonology*, *9*, 353-382.

Flege, J. E. (2003a). Methods for assessing the perception of vowels in a second language. *Issues in clinical linguistics*, 3-28.

Flege, J. E. (2003b). Assessing constraints on second-language segmental production and perception. *Phonetics and phonology in language comprehension and production: Differences and similarities*, 319-355.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, 233-277.

Flemming, E., 2004: “Contrast and perceptual distinctiveness” in *Phonetically Based Phonology*, Hayes, B., Kirchner, R., & Steriade, D. (Eds.), Cambridge University Press, pp. 232 – 272.

Husby, O. & Kløve, M. H. (2001):*Andrespråksfonologi – Teori og metodikk*, Gyldendal Akademiske Forlag, Oslo.

Iverson, P., Kuhl, P. K., Akahane-Yamada, R., Diesch, E., Tohkura, Y. I., Kettermann, A., & Siebert, C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. *Cognition*, *87*(1), B47-B57.

Kager, R. (1999). *Optimality Theory*, Cambridge University Press, Cambridge.

Katamba, F. (1989): *An Introduction to Phonology*, Longman Group, Essex.

Kenstowicz, M., & Suchato, A. (2006). Issues in loanword adaptation: A case study from Thai. *Lingua*, *116*(7), 921-949.

Kewley-Port, D. (1995): ”Thresholds for formant-frequency discrimination of vowels in consonantal context” in *J. Acoust. Soc. Am.* 97 (5), Pt. 1, May 1995, downloaded 22. Sep. 2012.

Kristoffersen, G. (2000): *The Phonology of Norwegian*, Oxford University Press, Oxford.

Kuhl, P. K. (2000): “A new view of language acquisition”, *PNAS* October 24, 2000 vol. 97 no. 22, pp. 11850-11857.

Kuhl, P.K. and Iverson, P. (1995): “Linguistic experience and the "perceptual magnet effect."”. In: Strange, W, (ed.) *Speech perception and linguistic experience: Issues in cross-language research*. (121 - 154). York Press: Baltimore, MD.

Kuhl, P. K. (1991): “Human adults and human infants show a "perceptual magnet effect" for the prototypes of speech categories, monkeys do not.” *Perception & Psychophysics*, 50, 93-107.

Ladefoged, P. (2003). *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. Malden, MA: Blackwell.

Lombardi, L. (2003). Second language data and constraints on Manner: explaining substitutions for the English interdentals. *Second Language Research*, *19*(3), 225-250.

McCarthy, J. J. (2002). *A thematic guide to Optimality Theory*. Cambridge University Press.

Maddieson, I., 1984: *Patterns of sounds*, Cambridge University Press, Cambridge.

Pätzold, M., & Simpson, A. P. (1997). Acoustic analysis of German vowels in the Kiel Corpus of Read Speech. *Arbeitsberichte des Instituts für Phonetik und digitale Spachverarbeitung Universität Kiel*, *32*, 215-247.

Prince, A., & Smolensky, P. (2008). *Optimality Theory: Constraint interaction in generative grammar*. Wiley. com.

Prince, A., & Smolensky, P. (1997). Optimality: From neural networks to universal grammar. *Science*, *275*(5306), 1604-1610.

Statistisk Sentralbyrå (2012): "Innvandring og innvandrere", http://www.ssb.no/innvandring/, retrieved 20. November 2012.

Van Dommelen, W.A. (p.c.): *Språkmøter*, personal correspondence, 2012.

Weiand, K. (2007). Implementing Escudero's model for the subset problem. *Rutgers Optimality Archive, 913*.

Zee, E., & Lee, W. S. (2001). An acoustical analysis of the vowels in beijing Mandarin. In *INTERSPEECH* (pp. 643-646).

Appendix 1 Information forms

1: For the Mandarin Chinese perception test informants

Cecilie Slinning Knudsen

Department of Language and Communication Studies, NTNU

Trondheim

cecilisl@stud.ntnu.no

Thank you for participating in my study!

This study is part of my Master's thesis in Phonology at NTNU.

Participation in this study includes the following:

1. A questionnaire with 7 questions about you, your age, your knowledge about languages in general and Norwegian in particular. None of these questions will gather sensitive information.

2. A perception test of some Norwegian and Mandarin sentences. This will last for about 30 minutes.

All the data will be anonymized and treated confidentially. After the data has been collected and analyzed, all questionnaires will be deleted. The project's completion is estimated to be at or around 10.07.2013.

This study is carried out under the supervision of dr.art. Jardar Eggesbø Abrahamsen, Department of Language and Communication Studies, NTNU. (jardar.abrahamsen@ntnu.no).

Participation is voluntary, and you may withdraw from the study at any time without being obliged to give a reason for this.

All questions concerning this study may be directed to me or my supervisor.

I hereby allow Cecilie Knudsen to use all the data she collects from me for her study,

-------------------------------------	--
Date, Place	Signature

2: For the Norwegian perception test informants

Cecilie Slinning Knudsen

Institutt for språk- og kommunikasjonsstudier, NTNU

Trondheim

cecilisl@stud.ntnu.no

Takk for at du vil delta i prosjektet mitt!

Denne lytteprøven er en del av masteroppgaven min i Fonologi ved NTNU.

Deltakelse i dette prosjektet innebærer en persepsjonstest av 100 norske setninger og vil vare ca. 15 minutter.

All data blir anonymisert og behandlet konfidensielt.

Dette prosjektet blir utført under veiledning av dr.art. Jardar Eggesbø Abrahamsen, Institutt for språk- og kommunikasjonsstudier, NTNU. (jardar.abrahamsen@ntnu.no).

Deltakelse er frivillig og du kan når som helst trekke deg fra prosjektet uten å gi en grunn for dette.

Alle spørsmål rundt prosjektet kan rettes til meg eller veilederen min.

Jeg tillater herved Cecilie Knudsen å bruke all data hun samler inn fra meg til prosjektet sitt,

-- -- -----

Dato, Sted Signatur

2 kopier: 1 til prosjektet og 1 til informanten

3: The Mandarin Chinese recording informant

Cecilie Knudsen

Department of Language and Communication Studies, NTNU

Trondheim

cecilisl@stud.ntnu.no

Thank you for participating in my study!

This study is part of my Master's thesis in Phonology at NTNU.

Participation in this study includes recording 3 sentences in your native language. The recording takes place in Fonlab, Department of Language and Communication Studies, NTNU, and will take between 20 and 30 minutes.

All the data will be anonymized and treated confidentially. The recordings will be used in a perception test after some digital manipulation. After the data of the perception tests have been collected and analyzed, all sound files will be deleted. The project's completion is estimated to be at or around 10.07.2013.

This study is carried out under the supervision of dr.art. Jardar Eggesbø Abrahamsen, Department of Language and Communication Studies, NTNU. (jardar.abrahamsen@ntnu.no).

Participation is voluntary, and you may withdraw from the study at any time without being obliged to give a reason for this.

All questions concerning this study may be directed to me or my supervisor.

I hereby allow Cecilie Knudsen to use all the data she collects from me for her study,

-------------------------------------	--
Date, Place	Signature

Appendix 2 Questionnaires

1: The questionnaire for the Mandarin Chinese informants (a formal questionnaire was not given to the Norwegian informants)

Questionnaire

Language:

Norwegian course level (1 or 3):

Given Name:

Age:

Phone number or e-mail:
(so that I can contact you when we're ready for the perception tests)

Are you a native speaker of Cantonese or Mandarin?

Where are you from? (City, area)

Are you, or have you, studied languages?

Which languages do you speak?

How long have you lived in Norway?

Have you lived in Norway before?

To what degree are you surrounded by Norwegian speech on a normal day?

Thank you for your participation!
Cecilie S. Knudsen

Appendix 3 Norwegian Social Science Data Services

Norsk samfunnsvitenskapelig datatjeneste AS
NORWEGIAN SOCIAL SCIENCE DATA SERVICES

NSD

Harald Hårfagres gate 29
N-5007 Bergen
Norway
Tel: +47-55 58 21 17
Fax: +47-55 58 96 50
nsd@nsd.uib.no
www.nsd.uib.no
Org.nr. 985 321 884

Jardar Eggesbø Abrahamsen
Institutt for språk- og kommunikasjonsstudier
NTNU
Dragvoll
7491 TRONDHEIM

Vår dato: 06.11.2012 Vår ref:31996 / 3 / MAS Deres dato: Deres ref:

TILBAKEMELDING PÅ MELDING OM BEHANDLING AV PERSONOPPLYSNINGER

Vi viser til melding om behandling av personopplysninger, mottatt 01.11.2012. Meldingen gjelder prosjektet:

31996	*Perception of Norwegian High Vowels*
Behandlingsansvarlig	*NTNU, ved institusjonens øverste leder*
Daglig ansvarlig	*Jardar Eggesbø Abrahamsen*
Student	*Cecilie Slinning Knudsen*

Personvernombudet har vurdert prosjektet og finner at behandlingen av personopplysninger er meldepliktig i henhold til personopplysningsloven § 31. Behandlingen tilfredsstiller kravene i personopplysningsloven.

Personvernombudets vurdering forutsetter at prosjektet gjennomføres i tråd med opplysningene gitt i meldeskjemaet, korrespondanse med ombudet, eventuelle kommentarer samt personopplysningsloven og helseregisterloven med forskrifter. Behandlingen av personopplysninger kan settes i gang.

Det gjøres oppmerksom på at det skal gis ny melding dersom behandlingen endres i forhold til de opplysninger som ligger til grunn for personvernombudets vurdering. Endringsmeldinger gis via et eget skjema, http://www.nsd.uib.no/personvern/forsk_stud/skjema.html. Det skal også gis melding etter tre år dersom prosjektet fortsatt pågår. Meldinger skal skje skriftlig til ombudet.

Personvernombudet har lagt ut opplysninger om prosjektet i en offentlig database, http://pvo.nsd.no/prosjekt.

Personvernombudet vil ved prosjektets avslutning, 10.07.2013, rette en henvendelse angående status for behandlingen av personopplysninger.

Vennlig hilsen

Vigdis Namtvedt Kvalheim

Mads Solberg

Mads Solberg tlf: 55 58 89 28
Vedlegg: Prosjektvurdering
Kopi: Cecilie Slinning Knudsen, Arne Bergsgårdsveg 20, Rom 11, 7033 TRONDHEIM

Avdelingskontorer / *District Offices*:
OSLO: NSD. Universitetet i Oslo, Postboks 1055 Blindern, 0316 Oslo. Tel: +47-22 85 52 11. nsd@uio.no
TRONDHEIM: NSD. Norges teknisk-naturvitenskapelige universitet, 7491 Trondheim. Tel: +47-73 59 19 07. kyrre.svarva@svt.ntnu.no
TROMSØ: NSD. SVF, Universitetet i Tromsø, 9037 Tromsø. Tel: +47-77 64 43 36. nsdmaa@sv.uit.no

Personvernombudet for forskning

Prosjektvurdering - Kommentar

Prosjektnr: 31996

Prosjektet skal redegjøre for fremmedspråkliges persepsjon av norske vokaler.

Ifølge prosjektmeldingen skal det innhentes skriftlig samtykke basert på muntlig og skriftlig informasjon om prosjektet og behandling av personopplysninger. Personvernombudet finner informasjonsskrivet tilfredsstillende utformet i henhold til personopplysningslovens vilkår forutsatt at en bemerkelse om at deltakelse er frivillig og at kandidaten kan trekke seg uten å oppgi grunn, samt dato for prosjektslutt (10.07.2013) tilføyes.

Innsamlede opplysninger registreres på privat pc. Personvernombudet legger til grunn at veileder og student setter seg inn i og etterfølger NTNU sine interne rutiner for datasikkerhet, spesielt med tanke på bruk av privat pc til oppbevaring av personidentifiserende data.

Prosjektet skal avsluttes 10.07.2013 og innsamlede opplysninger skal da anonymiseres og lydopptak slettes. Anonymisering innebærer at direkte personidentifiserende opplysninger som navn/koblingsnøkkel slettes, og at indirekte personidentifiserende opplysninger (sammenstilling av bakgrunnsopplysninger som f.eks. yrke, alder, kjønn) fjernes eller grovkategoriseres slik at ingen enkeltpersoner kan gjenkjennes i materialet.

Appendix 4 Resynthesis procedure

1. Isolate the vowel from the studio recording using ‘Audacity’: “Vowel” (“CY” – Chinese Y, “NY” – Norwegian Y)
2. ‘Praat’: “Vowel” *Convert* → *Resample*… New sampling frequency 10 000 Hz: “Vowel_Resampled”
3. “Vowel_Resampled” *Analyse Spectrum* → *To LPC(burg)*: “LPC”
4. *Analyse* → *Filter(inverse)* “Vowel_Resampled” and “LPC”: “Vowel_Source”
5. “Vowel_Resampled” *Analyse* → *To formant*: “Vowel_Resampled_Formant”
6. “Vowel_Resampled_Formant” *Convert* → *Down to FormantGrid*: “Vowel_Resampled_FormantGrid”
7. *View & Edit FormantGrid* → *Modify* → *Formula(frequencies)*: if row =2 then self -100 else self fi
8. “Vowel_Source” and “Vowel_Resampled_Formantgrid” *Filter*: “Vowel_Resampled_Source_Filt”
9. Rename with language, vowel and value codes and save (for example: CY_758 – Chinese “Y” 758 Hz)
10. Paste the manipulated vowel into the original sentence using ‘Audacity’
11. Adjust the dB of the vowel to agree with that of the original sentence using *Effekt* → *Forsterk*
12. Create a seamless transition between the signals by marking the transition area and apply *Effekt* → *Reparer*.
13. Rename with the prefix S for sentence (SCY_758).

Appendix 5 The perception test scripts

The L1 Norwegian perception test script:

"ooTextFile"

"Experiment x"

stimuliAreSounds? <yes>

stimulusFileNameHead = ""

stimulusFileNameTail = ".wav"

stimulusCarrierBefore = ""

stimulusCarrierAfter = ""

stimulusInitialSilenceDuration = 1.0 seconds

stimulusMedialSilenceDuration = 0

numberOfDifferentStimuli = 20

"SNY_0781" ""

"SNY_0899" ""

"SNY_1012" ""

"SNY_1120" ""

"SNY_1237" ""

"SNY_1324" ""

"SNY_1402" ""

"SNY_1518" ""

"SNY_1652" ""

"SNY_1793" ""

"SNY_1899" ""

"SNY_1956" ""

"SNY_2053" ""

"SNY_2122" ""

"SNY_2214" ""

"SNY_2335" ""

"SNY_2433" ""

"SNY_2518" ""

"SNY_2650" ""

"SNY_2754" ""

numberOfReplicationsPerStimulus = 5

breakAfterEvery = 10

randomize = <PermuteBalancedNoDoublets>

startText = "Klikk på musen for å starte"

runText = ""

pauseText = "Ei lita pause :) Klikk musen for å starte igjen"

endText = "Takk for at du deltok!"

maximumNumberOfReplays = 1

replayButton = 0.25 0.75 0.1 0.2 "Klikk her for å høre setningen på nytt" ""

okButton = 0.40 0.60 0.25 0.30 "OK" " "

oopsButton = 0 0 0 0 "" ""

responsesAreSounds? <no> "" "" "" "" 0 0

numberOfDifferentResponses = 3

0.25	0.75	0.75	0.85	"Ny" 30	""	"y"
0.25	0.75	0.55	0.65	"Nu" 30	""	"u"
0.25	0.75	0.35	0.45	"No" 30	""	"o"

numberOfGoodnessCategories = 0

randomize = <CyclicNonRandom>

The L1 Mandarin Chinese perception test script:

"ooTextFile"

"Experiment x"

stimuliAreSounds? <yes>

stimulusFileNameHead = ""

stimulusFileNameTail = ".wav"

stimulusCarrierBefore = ""

stimulusCarrierAfter = ""

stimulusInitialSilenceDuration = 1.0 seconds

stimulusMedialSilenceDuration = 0

numberOfDifferentStimuli = 19

"SCY_631" ""

"SCY_758" ""

"SCY_814" ""

"SCY_984" ""

"SCY_1033" ""

"SCY_1188" ""

"SCY_1262" ""

"SCY_1329" ""

"SCY_1412" ""

"SCY_1497" ""

"SCY_1600" ""

"SCY_1703" ""

"SCY_1861" ""

"SCY_1936" ""

"SCY_2074" ""

"SCY_2172" ""

"SCY_2259" ""

```
"SCY_2385" ""

"SCY_2547" ""



numberOfReplicationsPerStimulus = 5

breakAfterEvery = 0

randomize = <PermuteBalancedNoDoublets>

startText = "Click the mouse to start the test"

runText = ""

pauseText = "A little break :) Click the mouse to continue"

endText = "Thank you for your participation!"

maximumNumberOfReplays = 1

replayButton = 0.25 0.75 0.1 0.2 "Click here to listen to the sentence once more" ""

okButton = 0.40 0.60 0.25 0.30 "OK" " "

oopsButton = 0 0 0 0 "" ""

responsesAreSounds? <no> "" "" "" "" 0 0

numberOfDifferentResponses = 2

0.25    0.75    0.75    0.85    "ny" 30 ""    "y"

0.25    0.75    0.55    0.65    "no" 30 ""    "o"

numberOfGoodnessCategories = 0

# randomize = <CyclicNonRandom>
```

The L2 Norwegian perception test script:

"ooTextFile"

"Experiment x"

stimuliAreSounds? <yes>

stimulusFileNameHead = ""

stimulusFileNameTail = ".wav"

stimulusCarrierBefore = ""

stimulusCarrierAfter = ""

stimulusInitialSilenceDuration = 1.0 seconds

stimulusMedialSilenceDuration = 0

numberOfDifferentStimuli = 20

"SNY_0781" ""

"SNY_0899" ""

"SNY_1012" ""

"SNY_1120" ""

"SNY_1237" ""

"SNY_1324" ""

"SNY_1402" ""

"SNY_1518" ""

"SNY_1652" ""

"SNY_1793" ""

"SNY_1899" ""

"SNY_1956" ""

"SNY_2053" ""

"SNY_2122" ""

"SNY_2214" ""

"SNY_2335" ""

"SNY_2433" ""

```
"SNY_2518" ""
"SNY_2650" ""
"SNY_2754" ""

numberOfReplicationsPerStimulus = 5
breakAfterEvery = 10
randomize = <PermuteBalancedNoDoublets>
startText = "Click the mouse to start the test"
runText = ""
pauseText = "A little break :) Click the mouse to continue"
endText = "Thank you for your participation!"
maximumNumberOfReplays = 1
replayButton = 0.25 0.75 0.1 0.2 "Click here to listen to the sentence once more" ""
okButton = 0.40 0.60 0.25 0.30 "OK" " "
oopsButton = 0 0 0 0 "" ""
responsesAreSounds? <no> "" "" "" "" 0 0
numberOfDifferentResponses = 4

0.25    0.75    0.85    0.95    "Ni" 30     ""    "i"
0.25    0.75    0.65    0.75    "Ny" 30     ""    "y"
0.25    0.75    0.45    0.55    "Nu" 30     ""    "u"
0.25    0.75    0.25    0.35    "No" 30     ""    "o"
numberOfGoodnessCategories = 0
# randomize = <CyclicNonRandom>
```

Appendix 6 Norwegian results to the perception test

NM1			
F2	/u/	/ʉ/	/y/
781	5		
899	5		
1012	3	2	
1120		5	
1237	1	4	
1324		5	
1402		5	
1518		5	
1652		5	
1793		5	
1899		5	
1956		5	
2053		5	
2122		3	2
2214		1	4
2335			5
2433			5
2518			5
2650			5
2754			5

NM2			
F2	/u/	/ʉ/	/y/
781	5		
899	5		
1012	5		
1120	4	1	
1237	5		
1324	4	1	
1402	3	2	
1518	2	3	
1652		5	
1793		5	
1899		5	
1956		5	
2053		5	
2122		5	
2214			5
2335			5
2433			5
2518			5
2650			5
2754			5

NM3			
F2	/u/	/ʉ/	/y/
781	5		
899	5		
1012	1	4	
1120		5	
1237		5	
1324		5	
1402		5	
1518		5	
1652		5	
1793		5	
1899		5	
1956		5	
2053		5	
2122		5	
2214			5
2335			5
2433			5
2518			5
2650			5
2754			5

NM4			
F2	/u/	/ʉ/	/y/
781	5		
899	5		
1012	1	4	
1120		5	
1237		5	
1324		5	
1402		5	
1518		5	
1652		5	
1793		5	
1899		5	
1956		5	
2053		5	
2122		5	
2214			5
2335			5
2433			5
2518			5
2650			5
2754			5

NM5			
F2	/u/	/ʉ/	/y/
781	5		
899	5		
1012	5		
1120	4	1	
1237	3	2	
1324	2	3	
1402	1	4	
1518		5	
1652		5	
1793		5	
1899		5	
1956		5	
2053		4	1
2122		1	4
2214			5
2335			5
2433			5
2518			5
2650			5
2754			5 (-1 /i/)