The role of phonetics in phonological learning

Inaugural-Dissertation
zur Erlangung des
Doktorgrades der Philosophie (Dr. phil.)
durch die Philosophische Fakultät der
Heinrich-Heine-Universität Düsseldorf

vorgelegt von
Kim Strütjen

Betreuer:
Prof. Dr. Ruben van de Vijver
Prof. Dr. Dr. Peter Indefrey

eingereicht am 12. Januar 2021
Datum der Disputation: 01. September 2021

D61

Abstract

This dissertation investigates the role of different phonetic factors in phonological learning by considering theoretical and methodological aspects. On the theoretical side, my work deals with the phonetic factors and their interaction in learning. On the methodological side, it investigates the dependency on the type of test, the role of evidence for a morphophonological alternation during training and whether the learning or the generalization behavior is tested.

The learnability of vowel nasalization and retroflexion is studied with adult German native speakers in four artificial language learning experiments which are framed by typological surveys, perception and production experiments and an acoustic analysis.

The learning experiments on vowel nasalization show that perception biases the learning behavior more than articulation. They further demonstrate that generalizations are formed based on phonetic similarity. This is true in a perceptual as well as in a productive test. The learning experiments on retroflexion show that language-specific perceptual cues affect learning and generalizations more than language-general perceptual cues. However, without evidence for a morphophonological alternation, the learning performance becomes weaker, the influence of language-general perceptual cues increases and generalizations are impossible.

The results, thus, call for a tight integration of phonetics in phonology with phonological representations including phonetic details.

Abstract (Deutsch)

Diese Dissertation untersucht die Rolle verschiedener phonetischer Faktoren beim phonologischen Lernen unter Berücksichtigung theoretischer und methodischer Aspekte. Theoretisch befasst sie sich mit den phonetischen Faktoren und deren Interaktion beim Lernen. Methodisch testet sie, ob der Einfluss der Faktoren von der Testart und der Evidenz für eine morphophonologische Alternation im Training abhängig ist, und ob diese Faktoren das Lern- oder Generalisierungsverhalten betreffen.

Die Lernbarkeit von Vokalnasalierung und Retroflexion wird mit erwachsenen deutschen Muttersprachlern in vier Kunstsprachexperimenten untersucht – ergänzt um typologische Studien, Perzeptions- und Produktionsexperimente und eine akustische Analyse.

Die Lernexperimente zur Vokalnasalierung zeigen, dass die Perzeption das Lernen stärker beeinflusst als die Artikulation und dass Generalisierungen basierend auf phonetischer Ähnlichkeit gebildet werden – sowohl in einem perzeptiven als auch in einem produktiven Test. Die Lernexperimente zur Retroflexion zeigen, dass sprachspezifische perzeptuelle Cues das Lernen und Generalisieren stärker beeinflussen als sprachübergreifende. Ohne Evidenz für eine morphophonologische Alternation wird das Lernergebnis jedoch schwächer, steigt der Einfluss sprachübergreifender perzeptueller Cues und sind Generalisierungen nicht mehr möglich.

Somit sollten Phonetik und Phonologie eng miteinander verbunden sein und phonologische Repräsentationen phonetische Details enthalten.

Danksagung

Angefangen hat alles vor sieben Jahren mit zwei Masterseminaren bei Dinah Baer-Henney zum *Artificial Language Learning* und zu *Learning Biases*, aus denen nicht nur das Thema meiner Masterarbeit, sondern darauf aufbauend auch das Thema dieser Dissertation entstanden ist.

Ruben van de Vijver hat es mir anschließend ermöglicht, meine Begeisterung für experimentelles Arbeiten durch eine Anstellung als wissenschaftliche Mitarbeiterin am Institut für Sprache und Information in Düsseldorf beruflich weiterzuverfolgen. Neben der Lehre konnte ich Experimente planen, durchführen und analysieren und die Ergebnisse auf Konferenzen und in Kolloquien vorstellen. Ruben hat mir dazu immer konstruktives Feedback gegeben, stand mir mit Rat und Tat zur Seite, hat mir aber auch immer die Freiheit gelassen, das zu untersuchen, was ich wollte. Danke dafür, Ruben!

Auch Peter Indefreys Unterstützung bei der Planung und Auswertung der Experimente sowie seine Kommentare auf schriftliche Entwürfe haben einen großen und wichtigen Beitrag für das Gelingen dieser Arbeit geleistet. Dein anderer Blickwinkel als Psycholinguist war sehr hilfreich!

Ein besonderer Dank geht an die besten Kolleginnen der Welt: Dinah Baer-Henney, Jasmin Pfeifer und Jessica Nieder. Danke für eure Tipps und Tricks, euer Feedback, Anregungen für Experimente und natürlich die Ablenkung, wenn gerade mal alles zu viel wurde – sei es in Form von Kurzurlauben, Weihnachtsmarktbesuchen, ausgiebigem Essen oder einfach nur einer Tasse Kaffee. Ihr seid in den letzten Jahren zu echten Freundinnen geworden, die ich nicht missen möchte. Besonders hervorheben möchte ich an dieser Stelle Jessica, die immer ein offenes Ohr hat und ohne die ich weder das regelmäßige Pendeln mit der Bahn – das dank dir übrigens zu einer meiner Lieblingsbeschäftigungen geworden ist – noch das Projekt Promotion durchgehalten hätte.

Ein Dankeschön geht auch an alle anderen Kolleginnen und

Kollegen – aktuelle und ehemalige – am Institut, deren Namen zu nennen den Rahmen hier sprengen würde. Danke fürs Lieder raten, für Literaturtipps und technische Unterstützung, für gute Laune, offene Ohren und vieles mehr. Dank euch komme ich gerne ins Büro!

Lieben Dank an Christoph Draxler und Mareike Launer für die Programmierung der Online-Experimente, und natürlich an Erdin Mujezinovic, Isabel Schmidt und Mareike für die Hilfe bei der Datenerhebung, an Martin Rönsch und Yoshua Voigt für die Hilfe beim Annotieren und an Gabriella Gonçalves Pretelli und Marina Parigoridou für die Unterstützung bei der Stimulierstellung sowie an Alina Trauten und Lena Vennedey für ihre Vorstudien.

Ein großes Dankeschön an Gabriella, an Ramjiga Ramanathan sowie an zwei mir leider unbekannte Norweger, für die hier stellvertretend Martin Krämer steht. Danke für eure wunderbaren Stimmen und das Einsprechen der Stimuli.

Nicht zuletzt möchte ich meinen Probandinnen und Probanden danken, auch denen aus Vorstudien und denen, deren Daten in den Analysen leider nicht verwendet werden konnten. Danke für euer Engagement und eure Geduld sich Dutzende von *dobɛ̃ms* und *pɛɭɾas* anzuhören. Ohne euch wäre diese Arbeit nicht möglich gewesen!

Ein großer Dank geht auch an meine Studierenden fürs kritische Hinterfragen, Dinge aus einem anderen Blickwinkel sehen und für Diskussionen und neue Ideen auf Basis von Projektarbeiten.

Während der letzten Jahre hatte ich außerdem die Möglichkeit, meine Ergebnisse sowohl in diversen Kolloquien als auch auf verschiedenen internationalen Konferenzen vorzustellen. Danke an die Teilnehmerinnen und Teilnehmer der Phonetik/Phonologie- und Psycholinguistik-Kolloquien sowie der *P&Ps* in Marburg und München, der *OCP* in Budapest, der *PaPE* in Köln und der *LabPhon* in Lissabon für hilfreiches Feedback, aber auch für unterhaltsame Nachmittage und Abende. Gleiches gilt für die Teilnehmerinnen und Teilnehmer der *2nd Birmingham Statistics for Linguists Summer School* und natürlich an die wunderbaren Organisatoren – ihr habt mir gezeigt, welche Wunder R vollbringen kann.

Für ihre finanzielle Unterstützung in Form von Reisekostenzuschüssen und dem Frauenförderstipendium und der damit verbundenen Möglichkeit, diese Konferenzen zu besuchen, möchte ich mich außerdem bei den *Heine Research Academies* und der *philGRAD* sowie der Philosophischen Fakultät der Heinrich-Heine-Universität Düsseldorf bedanken.

Zum Schluss danke ich meiner Familie und meinen Freunden. Ihr habt mir immer den Rücken gestärkt und mir gezeigt, dass ich auf dem richtigen Weg bin. Mama und Papa, ihr habt niemals an mir gezweifelt, wart immer eine große Stütze für mich und habt mir all das ermöglicht. Danke, dass ihr immer für mich da seid! Marcel, du hast immer an mich geglaubt, jeden leisesten Zweifel sofort ausgeräumt und warst selbst aus der Ferne immer für mich da. Danke für deine Geduld, danke für alles.

Contents

List of Figures

List of Tables

List of Abbreviations

A1	amplitude of the highest harmonic
ant	anterior
C	consonant
C_1, C_2	first consonant, second consonant in an item
con	consonantal
cont	continuant
d’	dprime
dB	decibel
df	degrees of freedom
distr	distributed
DORS	dorsal
e.g.	example give
Est.	estimate
F0	fundamental frequency
F1, F2, F3	first, second, third formant
glmer	generalized linear mixed-effects model
H1, H2	first, second harmonic
HNR	harmonicity-to-noise-ratio
Hz	hertz
IPA	International Phonetic Association/Alphabet
kHz	kilohertz
L1, L2	first language (native language), second language
lm	linear model
lmer	linear mixed-effects model
MFC	multiple forced-choice
ms	milliseconds
N	nasal consonant
n	number
nas	nasal
p	p-value, conditional probability
P0, P1	amplitude of a nasal peak (harmonic) at about 250 Hz, 950 Hz
P-map	perceptual map

PHOIBLE	Phonetics Information Base and Lexicon
SD	standard deviation
SE	standard error
son	sonorant
t	t-value, Est. divided by SE in linear (mixed-effects) models
UPSID	UCLA Phonological Segment Inventory Database
V	oral vowel, volt
V_1, V_2	first vowel, second vowel in an item
Ṽ	nasal vowel
VOT	voice onset time
VSA	vowel space area
χ^2	chi-square-value, goodness of fit
z	z-value, Est. divided by SE in generalized linear mixed-effects models

1 Introduction

In this dissertation I will focus on the role of phonetics in phonological learning. Phonetics biases learners in the process of learning, which means that a certain phonological pattern is learned more easily than another phonological pattern (Baer-Henney, 2015; Hayes & Steriade, 2004; Peperkamp & Martin, 2016; White & Sundara, 2014; Wilson, 2006). This learning advantage is due to differences in phonetic details among phonological patterns, e.g. differences in the acoustic salience or differences in the articulatory or perceptual effort. Hence, learners use detailed phonetic knowledge to master the task of learning phonological patterns (Hayes, 1999).

By investigating the role of phonetics in phonological learning my work adds to the growing body of literature on learning biases. Most studies aimed at giving evidence for the existence of one of such learning biases, e.g. substantive bias (Baer-Henney, 2015; Peperkamp & Martin, 2016; White, 2017; Wilson, 2006), complexity bias (Moreton, 2008, 2012), sonority bias (Berent, Steriade, Lennertz & Vaknin, 2007; Gómez, Berent, Benavides-Varela, Bion, Cattarossi, Nespor & Mehler, 2014). I, however, will additionally explore how several learning biases interact with each other by testing the learnability of phonological patterns for which different learning biases make different predictions. In doing so, I will focus on biases grounded in phonetics – often called substantive biases (Baer-Henney, 2015; Wilson, 2006) or phonetic (naturalness) biases (Finley, 2012; Skoruppa, Lambrechts & Peperkamp, 2011). The term *substance* goes back to Chomsky (1965) and originally referred to the system of categories in the mental representations of our linguistic knowledge (Wilson, 2006). However, today *substance* refers to all aspects of grammar, which are grounded in the physical characteristics of speech. These are the articulatory, acoustic and perceptual characteristics of speech (Crystal, 2008). A similar definition was used by Seidl & Buckley (2005) who described *substance* as grounded in phonetics. They further put it on the same level as phonetic naturalness or unmarkedness. However, as Seidl & Buckley (2005) pointed out, the term *naturalness* is quite vague and the term *markedness* generally refers to frequency distributions so that substance should better be defined as phonetically grounded. As such, I will use the terms *phonetic bias* or *grounded in phonetics* as well. I will further examine the influence of phonetic biases that are based on ease of articulation or ease of perception and the influence of phonetic similarity on the learning behavior and on the generalization behavior. The learning behavior refers to the participants' performance in the trained pattern only, whereas the generalization behavior refers to the application of the trained pattern to novel patterns.

I will investigate these phonetic factors with German native speakers in four artificial language learning experiments with vowel nasalization and retroflexion. Vowel nasalization and retroflexion are both phonological patterns, which are not used phonemically in German – retroflexion is even not used allophonically in German (Wiese, 1996). In the artificial language learning experiments testing vowel nasalization I will explore whether ease of articulation, ease of perception or phonetic similarity have a stronger effect on learning a phonological pattern and generalizing it to novel patterns. In addition to that, I will test whether the effects depend on the type of test – production or perception. In the learning experiments testing retroflexion I will investigate whether ease of perception or the perceptual similarity to phonemes of the native language (L1) have a stronger effect on learning a phonological pattern and generalizing it to novel patterns. Besides, I will test whether the effects depend on whether direct evidence for a morphophonological alternation is provided during training or not. By not providing evidence for an alternation I will further investigate whether a phonological pattern is actually generalizable due to pure transfer (as proposed by e.g. Berent (2013) and Berent, Wilson, Marcus & Bemis (2012)) or whether learners need to know about the relation between morphophonological forms to accurately generalize a phonological pattern. Short typological surveys as well as production and perception experiments frame these artificial language learning experiments to establish whether what is assumed to be articulatory and perceptually easy according to the literature is indeed articulatory and perceptually easy for German native speakers.

1.1 Phonetic effects on learning phonological patterns

Do learners bring phonetic knowledge to the task of learning phonological alternations? When we generalize from observed phonological patterns to unfamiliar phonological patterns we have not observed, what is the role of the phonetics of the observed patterns and of the novel patterns? Do phonetic similarities between sound patterns play a role or are generalizations mainly determined by ease of articulation or ease of perception?

Testing the learning behavior – that is how well participants perform when being tested on the same pattern that they have been trained on, e.g. nasalization of [i] – gives insights into what mechanisms are used during learning and shows whether learners bring phonetic knowledge to the task of phonological learning. However, the process of learning involves not only memorizing a certain pattern but incorporates also how this newly gained knowledge is applied to other – untrained – patterns, e.g. nasalization of [ɛ] and [a]. I will refer to this extension of a trained pattern to a novel one as *generalizations* in this dissertation. With the term *learning* I will refer to the 'pure' learning of exactly that pattern that has been trained – to contrast it with *generalizations*. For example, in the learning experiment on vowel nasalization (see chapter 2.4) one group of learners will be trained on high nasal vowels, but at test this group will be confronted not only with high nasal vowels (trained) to test their learning behavior but also with mid and low nasal vowels (untrained) to test their generalization behavior. Whenever learning (behavior) is not used to contrast with generalization (behavior), *learning* will refer to the broader mental process of acquiring something new as used in common language use, e.g. as in terms like ‘phonological learning’. This wording is in line with several other experimental studies on phonological learning, e.g. Baer-Henney & van de Vijver (2012), Becker, Nevins & Levine (2012), Cristià, Mielke, Daland & Peperkamp (2013), Finley & Badecker (2009), Linzen & Gallagher (2017) and Wilson (2006).

The above-mentioned questions can be answered by looking at how well participants learn phonological patterns which differ in ease of articulation, ease of perception and phonetic similarity and how they generalize the learned pattern to novel, similar patterns.

Ease of articulation refers to how easy it is to articulate a certain sound or pattern. Something is easy to articulate when it requires e.g. less muscular activity or small changes in vocal tract configurations (Hayes & Steriade, 2004). Let me illustrate this using the so-called alternation between the [iç]- and the [ax]-*Laut* in German. In German the palatal fricative [ç] is articulated after front vowels, e.g. [i], whereas the velar fricative [x] is articulated after back vowels, e.g. [a] (Wiese, 1996). The reason for this lies in the places of articulation of the vowels and the fricatives: Depending on the place of articulation of the preceding vowel the place of articulation of the fricative alternates to shorten the articulatory distance between vowel and fricative, hence avoiding large movements of the tongue during the articulation of this sound sequence. Producing [ix] or [aç] would be more difficult due to the articulatory distance between the vowels and the fricatives. One further example is the process of velar palatalization by which velar plosives, e.g. [k], become palatoalveolar affricates, e.g. [t͡ʃ]. Velar palatalization is easier before front vowels, e.g. [i], than before central or back vowels, e.g. [e] or [a] (Wilson, 2006). This is because due to the velar palatalization the place of articulation of the plosive is fronted from the velar region to the palatal region. This means that the places of articulation of the consonant and the vowel are more equal to each other in [t͡ʃi] than in [t͡ʃe] or [t͡ʃa], and hence their articulation requires less muscular activity of the tongue – similar to the alternation between [iç] and [ax] in German described above. When speakers rely on ease of articulation, those participants who are trained on the pattern which is easiest to articulate – that is a pattern which is articulated with less muscular activity or smaller changes in vocal tract configurations, e.g. [t͡ʃi], compared to another one – should be more successful in learning than those who are trained on a pattern which is more difficult to articulate – that is a pattern which is articulated with more muscular activity or greater changes in vocal tract configurations, e.g. [t͡ʃe] or [t͡ʃa], compared to another one. When speakers base their generalizations on ease of articulation, they should generalize more to a pattern that is easy to articulate, e.g. [t͡ʃi]. This is exactly what Wilson (2006) found in his learning experiment. For details and further experimental studies testing a learning bias based on ease of articulation see chapter 1.1.1.

In Figure 1 you can see how ease of articulation might affect phonological learning in an experimental setting using the example of velar palatalization described above. Note that for illustrative purposes it is assumed that ease of articulation is the only factor affecting learning and generalizations. In this Figure A', B' and C' are sound patterns which differ in how easy they are to articulate. Let us assume that A' is easier to articulate than B' and C'. This is illustrated with the thick rectangle around A' as compared to the thin rectangles around B' and C'. Imagine an experimental setting in which three groups of participants are trained on the sound patterns A', B' and C' by means of auditory exposure, whereby each group is trained on only one of these patterns. As A' is easiest to articulate, it will also be easiest to learn. Learning is illustrated with the arrows connecting the sound patterns in the rectangles and the learned categories in the circles. The thick arrow connecting A' and A as well as the thick circle around A mean better learning than the thin arrows connecting B' or C' with B or C as well as the thin circles around B and C. The arrows connecting the learned categories A, B and C with each other illustrate the generalizations. As can be seen, it is easier to generalize to the articulatory easy A (thick arrows) than to the articulatory more difficult categories B and C (thin arrows).

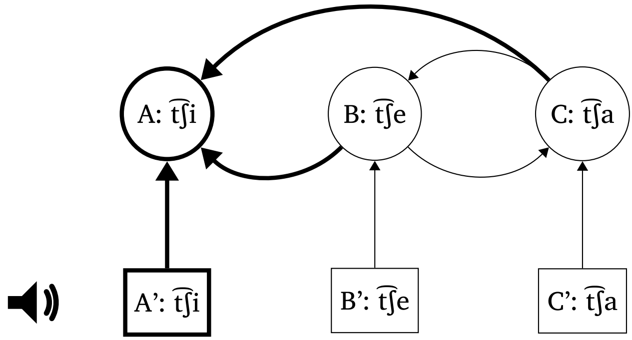

Figure 1: Learning and generalizing based on ease of articulation using the example of velar palatalization.

Ease of perception refers to how easy it is to perceive a certain sound or pattern. Something is easy to perceive when there are no perceptually similar sounds in the inventory or in the context under investigation (Hayes & Steriade, 2004). Let me illustrate this with the perception of non-native vowels: It is easier to perceive the contrast between two non-native vowels, which are phonetically more different from each other, than the contrast between two non-native vowels, which are phonetically more similar to each other. For German native speakers who are not familiar with the English vowels [ʌ] as in *but*, [ɑ] as in *palm* and [ɜ] as in *bird* it is more difficult to perceive the contrast between [ʌ] and [ɑ] than to perceive the contrast between [ɑ] and [ɜ] as I found out in a short survey. This is because [ʌ] is a back open-mid vowel, whereas [ɑ] is a back open vowel – hence both vowels differ only in their height but not in their backness. [ɜ], however, is a central open-mid vowel, which means that it differs from [ɑ] not only in height but also in backness. Thus, the less phonetically similar two sounds are, the easier it is to perceive their contrast. A further example can be found in the perception of rounding harmony – a process in which all vowels in a domain are either round, e.g. [u] and [o], or not, e.g. [i] and [e], – which is easier with mid vowels than with high vowels (Kaun, 2004). The reason for this lies in the articulation of these vowels: All high vowels – independent of their rounding – are produced with more lip rounding than mid vowels, which makes high unround vowels sound even a bit round. When speakers rely on ease of perception, those participants who are trained on the pattern which is easiest to perceive – that is a pattern which forms a larger contrast, e.g. rounding harmony of mid vowel, than another one – should be more successful in learning than those who are trained on a pattern which is more difficult to perceive – that is a pattern which forms a smaller contrast, e.g. rounding harmony of high vowels, than another one. When speakers base their generalizations on ease of perception, they should generalize more to a pattern that is easy to perceive, e.g. rounding harmony of mid vowels. In her study, Finley (2012) found that learners exposed to mid rounding harmony indeed generalized their learned pattern to novel items with mid vowels, whereas learners exposed to high rounding harmony did not generalize their learned pattern to novel items with high vowels. Unfortunately, no generalizations to other vowel heights were tested. For details and further experimental studies testing a learning bias based on ease of perception see chapter 1.1.1.

In Figure 2 you can see how ease of perception might affect phonological learning in an experimental setting. Note that for illustrative purposes it is assumed that ease of perception is the only factor affecting learning and generalizations. In this Figure A’, B’ and C’ are sound patterns which differ in how easy they are to perceive. Let us assume that A’, e.g. rounding harmony with mid vowels, is easier to perceive than B’, e.g. rounding harmony with high vowels, and C’. This is illustrated with the thick rectangle around A’ as compared to the thin rectangle around B’ and C’. Imagine an experimental setting in which three groups of participants are trained on the sound patterns A’, B’ and C’ by means of auditory exposure, whereby each group is trained on one pattern only. As A’ is easiest to perceive, it is also easiest to learn. Learning is illustrated with the arrows connecting the sound patterns in the rectangles and the learned categories in the circles. The thick arrow connecting A’ and A as well as the thick circle around A mean better learning than the thin arrows connecting B’ or C’ with B or C as well as the thin circles around B and C. The arrows connecting the learned categories A, B and C with each other illustrate the generalizations. As can be seen, it is easier to generalize to the perceptually easy A (thick arrows), e.g. rounding harmony with mid vowels, than to the perceptually more difficult categories B and C (thin arrows), e.g. rounding harmony with high vowels.

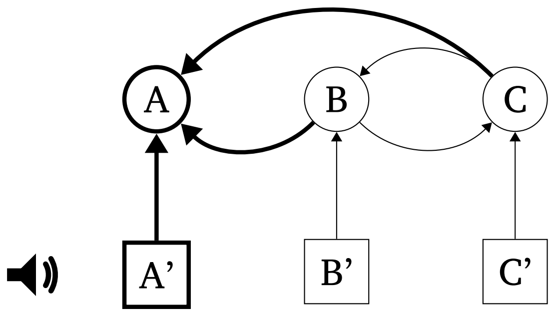

Figure 2: Learning and generalizing based on ease of perception.

Note that Figure 1 (learning and generalizing based on ease of articulation) and Figure 2 (learning and generalizing based on ease of perception) do not differ in their structure, as the mechanisms applying in learning and generalizing are the same. The only difference is the reason why a certain pattern – in this case A’ – is easy to learn and consequently easy to generalize to –

either due to articulatory ease or due to perceptual ease.

Phonetic similarity refers to how similar certain sounds or patterns are. Sounds or patterns are phonetically similar to each other when they have similar phonetic features or when they are acoustically and perceptually close to each other (Steriade, 2001b, 2009). Phonetic similarity acts on two different levels. Either it refers to the similarity between the learned pattern and the generalized pattern or it refers to the similarity between a non-native sound in the learned pattern and a native sound. There is also a third level on which phonetic similarity can act: the similarity between alternating sounds. However, this similarity resembles what I refer to as ease of perception that deals with the perceptibility of contrasting or alternating sounds. Thus, I will not focus on it here. I will now briefly explain the difference between these forms of phonetic similarity.

First, the learned pattern (the trained pattern) and the generalized pattern (the untrained pattern) can be more or less phonetically similar to each other. In the following I will refer to this kind of similarity as phonetic similarity. It can be illustrated with the vowels [a], [ɛ] and [i]. Low [a] and mid [ɛ] are phonetically more similar to each other than mid [ɛ] and high [i]. The reason for this is that [ɛ] is a lower mid vowel, with formant frequencies closer to low [a] than to high [i] (see chapter 1.3 for details). The same is true for the nasal vowels: [ɛ̃] is phonetically more similar to [ã] than to [ĩ]. When speakers base their generalizations on phonetic similarity, they should generalize more to a pattern that is phonetically similar – measured in terms of phonetic features or perceptual and acoustic characteristics between two or more sounds or patterns – to the learned one. I will investigate this in chapter 2.4. Note that phonetic similarity makes no predictions regarding learning – it does not predict a learning advantage for one pattern over the other one as e.g. ease of perception or ease of articulation do – as it refers to the similarity between the learned pattern and the generalized pattern only.

Second, a non-native sound in a pattern that should be learned can be phonetically similar to a native sound. In the following I will refer to this kind of similarity as perceptual similarity to native language phonemes (Kuhl, 1991; Kuhl & Iverson, 1995). It can be illustrated with the retroflex fricative [ʂ] that is phonetically more similar to the postalveolar fricative [ʃ] than to the alveolar fricative [s] for German native speakers (see chapter 4.3). When speakers rely on the perceptual similarity to native language phonemes – measured in terms of phonetic features or perceptual and acoustic characteristics between two or more sounds or patterns –, those participants who are trained on the pattern that is phonetically most similar to an already known L1-pattern should be more successful in learning than those who are trained on a pattern that is less similar to an already known L1-pattern. When speakers base their generalizations on perceptual similarity to native language phonemes, they should generalize less to a pattern that is similar to an L1-pattern, as this L1-pattern does not involve the trained pattern but an incorrect mapping onto a native pattern. I will investigate this in chapter 4.4.

In this dissertation I will focus on both aspects of phonetic similarity in learning and generalizing phonological alternations. As the second aspect is affected by the native language inventory there will be a separate chapter on the similarity to native sounds below. More information on phonetic similarity and experimental studies testing a learning bias based on phonetic similarity can be found in chapters 1.1.2-1.1.3.

In Figure 3 you can see how phonetic similarity might affect phonological generalizations in an experimental setting. Note that for illustrative purposes it is assumed that phonetic similarity is the only factor affecting phonological generalizations. In this Figure A', B' and C' are sound patterns which differ in their phonetic similarity. B', e.g. [ɛ], is phonetically more similar to A', e.g. [a], than to C', e.g. [i]. This is illustrated with the distance between A', B' and C'. In this case neither of A', B' or C' is easier to learn so that the arrows connecting X' and X as well as the rectangles around A', B' and C' have all the same thickness. However, the generalizations differ as is illustrated with the arrows connecting the learned categories A, B and C. As can be seen, it is easier to generalize from B to A (thick arrows), and vice versa, than from B to C (thin arrows), and vice versa. The most difficult generalizations are from A to C (dotted arrows), and vice versa.

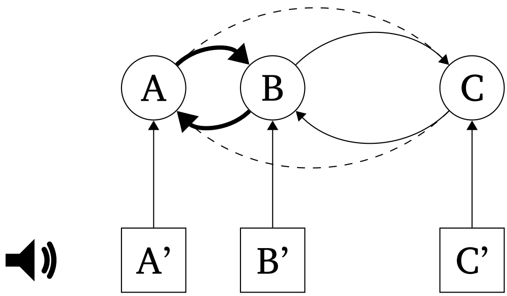

Figure 3: Generalizations based on phonetic similarity.

In Figure 4 you can see how the perceptual similarity to a native language phoneme might affect phonological learning and generalizations in an experimental setting. Note that for illustrative purposes it is assumed that the perceptual similarity to a native language phoneme is the only factor affecting phonological learning and generalizations. In this Figure A’, B’ and C’ are sound patterns which differ in their perceptual similarity to the native language phoneme A”, B” and C”. Only the native category C” is perceptually very similar to the trained category C so that it is (mis)learned as C”. This is illustrated with the thick arrow connecting C’ with C” as well as with the overlapping circles of C and C”. The native categories A” and B”, however, are more distant from the learned categories A and B so that they do not affect learning. This is illustrated with the non-overlapping circles of A and A” as well as of B and B”. Learning a native category is easier than learning a non-native category, which is illustrated with the thick and thin arrows connecting the rectangles with the circles as well as with the thickness of the circles around the learned categories. Generalizations from and to the mislearned category C” are not possible as C” does not have the same phonetic and phonological characteristics as A and B. Thus, only generalizations from and to A and B are possible.

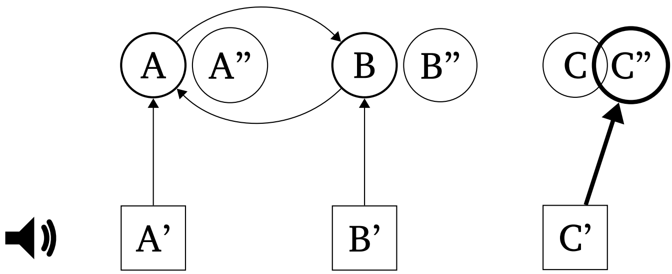

Figure 4: Learning and generalizing based on perceptual similarity to L1.

In the following chapter 1.1.1 I will describe each of the phonetic factors under investigation in detail and I will provide evidence for their role in phonological learning. After that I will continue with a description of the method used in the present experimental studies – artificial language learning experiments – in chapter 1.2 and an explanation why vowel nasalization (chapter 1.3) and retroflexion (chapter 1.4) are appropriate sound patterns to explore the interaction of phonetic biases during phonological learning. The present chapter ends with an overview of my experimental studies (chapter 1.5). The subsequent chapters focus on the four experimental studies on vowel nasalization (chapters 2 and 3) and retroflexion (chapters 4 and 5) whose results are summarized and discussed in chapter 6. A final conclusion is drawn in chapter 7.

1.1.1 Phonetically based phonology: ease of articulation and ease of perception

According to the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006) there is a learning bias for phonological patterns that results from ease of articulation or ease of perception. This means that phonological patterns that are either easier to articulate – due to e.g. less muscular activity – or easier to perceive – due to e.g. less perceptually overlap – than other phonological patterns, or both, are easier to learn.

Within the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006) it is proposed that knowledge of the phonetic substance of a phonological pattern – its ease of articulation and perception – is part of the learners' linguistic knowledge and that learners are predisposed towards patterns motivated in this way. The predisposition is encoded in the learners' grammars as markedness that reveals itself in typological distributions (Archangeli & Pulleyblank, 1994; Finley, 2012; Hayes & Steriade, 2004; Wilson, 2006). Crudely stated, unmarked patterns are easy to articulate or easy to distinguish, or both, while marked patterns are difficult to articulate or distinguish. Phonetic substance acts as a learning preference so that learning an unmarked pattern is easier than learning a marked pattern. As a consequence of this learning preference, unmarked patterns are attested more abundantly cross-linguistically than marked patterns. In addition to that, children prefer unmarked patterns over marked patterns (Demuth, 1995; Jusczyk, Smolensky & Allocco, 2002). An example of this is final devoicing which is unmarked in comparison to final voicing. It is perceptually difficult to distinguish voiced and voiceless obstruents word-finally (Keating, 1984; Steriade, 1997), and this knowledge is encoded as markedness and reflected typologically. There are many unrelated languages in which the voicing contrast among obstruents is almost completely neutralized in word-final position (in Indo-European languages such as Dutch (Booij, 1995; Warner, Jongman, Sereno & Kemps, 2004), German (Röttger, Winter, Grawunder, Kirby & Grice, 2014) and Polish (Slowiaczek & Dinnsen, 1985), in the Semitic language Maltese (Borg & Azzopardi-Alexander, 1997) and in the Altaic language Turkish (Kopkallı, 1993)), whereas there are only very few languages (e.g. the Nakh-Daghestanian language Lezgian) which have a process in which obstruents at the end of a word are voiced (Yu, 2004). The abundance of final devoicing and the scarcity of the opposite pattern – final voicing – are hypothesized by phonetically based phonology to be a consequence of learning (Hayes & Steriade, 2004; Wilson, 2006): There is a learning preference for phonetically well motivated patterns in comparison to phonetically less motivated patterns.

To test the existence of a learning preference for unmarked patterns over marked patterns or the existence of a learning preference for articulatory or perceptually easy patterns over articulatory or perceptually difficult patterns, participants are tested on how well they can learn a specific phonological pattern and how well they can generalize this pattern to novel instances in an experimental set-up (for details on this method see chapter 1.2).

Wilson (2006) investigated whether phonetic substance, instantiated as ease of articulation or ease of perception, acts as a bias in learning phonological patterns. As an empirical basis he used velar palatalization, a process in which velar stops ([k] and [g]) are pronounced as palatoalveolar affricates ([t͡ʃ] and [d͡ʒ]) before front vowels ([i] and [e]). If a language palatalizes velars before [e], it also palatalizes them before [i]. This is the result of a cascade of phonetic factors that starts out with articulation. The more front a vowel is before a velar stop, the more front its place of articulation is. As a result the velar receives a place of articulation very similar to the place of articulation of a palatoalveolar (Keating & Lahiri, 1993). A consequence of this fronting is that the noise spectrum of the release burst is acoustically similar to the noise spectrum of a palatoalveolar affricate, which makes them perceptually similar (Guion, 1996, 1998;

Keating & Lahiri, 1993). In other words, [k] is more like [t͡ʃ] before [i] than before [e], since [i] is more front than [e], and [k] is more like [t͡ʃ] before [e] than before [a], since [e] is more front than [a]. These facts reflect the typological distribution of velar palatalization as well (Guion, 1996, 1998). In sound patterns in languages of the world velar palatalization occurs more frequently before front vowels than before back vowels, and if it occurs before back vowels, it also occurs before front vowels (Guion, 1996, 1998). Wilson (2006) investigated learning of velar palatalization by means of a language game. Participants were taught a small discourse, in which one group first heard "I say [ki]" followed by "You say..." after which they would have to respond with "[t͡ʃi]". Another group was taught that after they heard [ke] they would have to respond with [t͡ʃe]. After training both groups of participants were prompted to respond to [ki]- and [ke]-items. Wilson (2006) found that participants trained with [ke]-items generalized velar palatalization to [ki]-items, thus responding with [t͡ʃe] and [t͡ʃi]. This is due to the fact that [i] is phonetically more front than [e], and if [e] triggers palatalization, then [i] should trigger it as well. This is also the pattern found in typology: If a language has palatalization before [e], it will also have palatalization before [i]. In contrast, participants who were trained with [ki]-items did not generalize palatalization to [ke]-items, thus only responding with [t͡ʃi], but not with [t͡ʃe]. This is due to the fact that [i] is more front than [e], and if [i] triggers palatalization, this does not mean that [e] should also trigger it. This pattern, too, is found in typology: If a language has palatalization before [i], it will not necessarily have palatalization before [e]. Wilson (2006) interpreted his results as support for the theory that typological patterns are the result of a preference in learning for phonetically based phonological patterns. Similar findings and conclusions concerning the role of ease of articulation and ease of perception are reported in Baer-Henney & van de Vijver (2012), Baer-Henney, Kügler & van de Vijver (2015a), Finley (2012, 2017) and Finley & Badecker (2009).

Finley (2012) investigated the learning of rounding harmony in an artificial language. Cross-linguistically mid vowels, e.g. [e] and [o], are more often triggers – that are sounds which initiate a phonological alternation – of rounding harmony than high vowels, e.g. [i] and [u], because mid round and unround vowels are easier to distinguish from each other than high round and unround vowels; high vowels, even unround ones, always sound somewhat round (Kaun, 2004). Finley (2012) created an artificial language that was taught to native speakers of English, a language without vowel harmony. One group of learners was taught a rounding harmony pattern involving mid vowels and another group a rounding harmony pattern involving high vowels. If ease of perception affects learning, the grammar with mid vowels as triggers of harmony should be easier to learn than the grammar with high vowels as triggers. This is indeed what Finley (2012) found. She concluded that this finding supports the existence of a learning preference that is independent of the native language of a learner and that this learning preference is therefore also likely responsible for typology.

Martin & Peperkamp (2020) found a learning preference for phonetically motivated patterns that are typologically frequent in their experiment as well. They tested the learnability of a vowel backness harmony (phonetically motivated and typologically frequent) and a vowel backness disharmony (not phonetically motivated and typologically infrequent) with native speakers of American English. Those participants who had been exposed to the harmonic pattern performed better at test than those participants who had been exposed to the disharmonic pattern. Martin & Peperkamp (2020) thus concluded that phonetic substance acts as a learning bias leading to a learning advantage for harmonic patterns over disharmonic patterns. This is supported by the findings of Martin & White (2021).

There are, however, also experimental studies that do not fully support a learning bias based on phonetics. Glewwe (2019) focused in her work on biases in phonotactic learning – learning positional distributions of phonological alternations – and in alternation learning – learning phonological alternations independent of its position in the word. In the first set of experiments she investigated whether English native speakers were better in learning a phonetically motivated phonotactic pattern that is typologically unmarked and whether they extended it based on the implicational universals from typology. In doing so she focused on the positional distribution of place contrasts and of voicing. According to Steriade (2001a) consonants are easier to distinguish before a vowel than after a vowel. Moreover, obstruent voicing is easier to perceive word-initially than word-finally – similar to the place contrasts. Consequently, if place of articulation contrasts after a vowel – that is the perceptually difficult condition –, it should also contrast before a vowel – that is the perceptually easy condition –, and if voicing contrasts word-finally – that is the perceptually difficult condition –, it should also contrast word-initially – that is the perceptually easy condition. This is exactly what is found in the languages of the world. However, in Glewwe (2019)'s experiments the participants did not extend the place contrast more from postvocalic position to prevocalic position than from prevocalic to postvocalic position – showing no support for a phonetic bias in phonotactic learning. Concerning the voicing contrast, there was some support for a phonetic bias, as some participants tended to extend the voicing contrast more from word-final position to word-initial position than from word-initial position to word-final position; other participants, however, did not. Glewwe (2019) further found that those participants who learned that voiceless obstruents are not allowed performed better than those participants who learned that voiced obstruents are not allowed. She explained these results with the structure of the stimuli: Except for the obstruent in word-final or word-initial position only voiced sonorants were used as consonants. Thus, those participants who learned that voiceless obstruents are not allowed word-finally or word-initially, had to learn only a one-feature constraint (#*[−voice] or *[−voice]#), whereas participants who learned that voiced obstruents were not allowed had to learn a two-feature constraint (#*[−son, +voice] or *[−son, +voice]#). This two-feature constraint is more complex than the one-feature constraint, which means that the two-feature constraint is also more difficult to learn than the one-feature constraint. The effect of the structural complexity of a pattern on its learnability is manifested in a complexity bias or a simplicity bias[1]: Learners are biased to prefer structurally simple patterns or constraints (Moreton, 2008, 2012) (for details see Glewwe (2019): chapter 3).

In the second set of experiments Glewwe (2019) investigated whether English native speakers were better in learning a phonetically motivated alternation which is typological unmarked. Participants were exposed either to a final voicing pattern or to a final devoicing pattern, whereas final devoicing is phonetically motivated as voiceless plosives are easier to produce word-finally than voiced plosives (Westbury & Keating, 1986). Consequently, final devoicing should be learned better than final voicing. The experimental results, however, showed the opposite pattern: Final voicing was learned better than final devoicing. In this set of experiments Glewwe (2019) argued again that a simplicity bias might explain the results. Thus, these experiments investigated the learnability of phonotactic patterns and its implicational universals as well as the learnability of phonological alternations, all of which were grounded either in perceptual ease or in articulatory ease. Glewwe (2019) interpreted some of her results in favor of a phonetic bias – others showing evidence for a simplicity bias – and argued for a robust simplicity bias as well as for a distinction between perceptually grounded and articulatory grounded phonetic biases. She further claimed that only a perceptually grounded phonetic bias influences phonological learning. This is because Glewwe (2019) found a generalization advantage based on perceptual ease (positional distribution of voicing) but no learning advantage based on articulatory ease (final voicing vs. final devoicing).

[1] The terms *simplicity bias* and *complexity bias* are used interchangeably. I decided to stick with the term *simplicity bias* as learners are biased to prefer simple patterns – similar to a phonetic bias due to which learners are biased to prefer the phonetically motivated pattern.

In Glewwe (2019)'s study a phonetic learning bias did not show up in all experiments. There are, in addition to that, further reasons to assume that it is not clear whether the above-mentioned results of previous studies support the claim that the ease of articulation or the ease of perception of a phonological pattern afford it a learning advantage, and that this learning advantage is reflected in the typological distribution of a pattern. The existence of such a learning bias based on phonetics is thus a contentious matter. See Moreton & Pater (2012a,b) for a review.

The study of Kimper (2016), for example, was similar to the one of Finley (2012), who found a learning advantage for the perceptually easy pattern in her experiment on vowel harmony (see above). Kimper (2016) found evidence for a phonetic bias, too. However, this bias did only show up in some of the learners. Moreover, training included a verification phase with feedback so that learning was explicit and not implicit as it should be in artificial language learning experiments (see chapter 1.2). The results of Baer-Henney (2015)'s experiment which she described in chapter 4 are problematic as well. She trained native German participants on intervocalic voicing of different consonants to test for the gradience of a phonetic bias. Gradience refers to a continuous scale without clear boundaries between two or more categories (Crystal, 2008). The larger the oral cavity, the easier it is to maintain voicing (Westbury & Keating, 1986). Thus, intervocalic voicing of labial consonants is articulatory easier than intervocalic voicing of coronal consonants or dorsal consonants, meaning a learning advantage in terms of a phonetic bias for labial consonants or a learning disadvantage for dorsal consonants. She found indeed a learning disadvantage for dorsal consonants. For the generalizations one would predict more generalizations from dorsal to other places and fewer generalizations from labial to other places. However, this was not fulfilled. Participants trained on labials generalized most. Thus, not every aspect of a phonetic bias could be given evidence for. Moreover, German native speakers might have experience with intervocalic voicing as German has final devoicing: Conversely, final devoicing can be interpreted as a process of intervocalic voicing in some contexts, e.g. [ta:k] *Tag* 'day' – [taːgə] *Tage* 'days'. Baer-Henney (2015)'s experimental results may thus be influenced by the occurrences of the consonants under investigation in German according to Glewwe (2019).

Another problem with studies investigating a phonetic bias is that phonetic naturalness is often confounded with simplicity, thus giving evidence for a simplicity bias instead. Peperkamp, Skoruppa & Dupoux (2006) trained participants either on the phonetically natural rule of intervocalic voicing or on a phonetically unnatural rule, in which /g/ surfaces as [f] intervocalically. The results showed a learning advantage for the phonetically natural pattern over the phonetically unnatural pattern so that the authors claimed to have found evidence for a phonetic bias. However, they confounded phonetic naturalness with structural simplicity, as the voicing alternation can be described by a one-feature change from [−voiced] to [+voiced] whereas the phonetically unnatural alternation cannot.

Another problem with investigating the role of phonetic biases in learning is that many experiments reached null results. Pycha, Nowak, Shin & Shosted (2003) compared the learning of vowel harmony – a phonetically motivated and unmarked pattern – and vowel disharmony. Both patterns could be predicted by only one feature: backness. They found no difference in learning. However, they gave further support for a simplicity bias, as the learning of a third pattern, an arbitrary rule involving more than one feature change, was learned worse than the vowel harmony and the vowel disharmony pattern. Testing the same with a rounding harmony, Skoruppa & Peperkamp (2011) found no difference in learning a harmony pattern and in learning a disharmony pattern either. However, again, a mixed-harmony/disharmony-pattern involving more than one feature change was learned worst.[2]

The results of the study of Wilson (2006), described in detail

[2]Note that Martin & Peperkamp (2020) found a learning advantage for a vowel harmony pattern over a vowel disharmony pattern without confounding phonetic naturalness with structural simplicity.

above, have been criticized as well (Moreton & Pater, 2012b). Even though the results Wilson (2006) obtained in his language game for learning palatalization of voiceless velar stops are congruent with the typological distribution of velar palatalization, the results concerning voiced velar stops are not. This is problematic since the phonetic effects that drive the typological distribution of palatalization of voiceless velars are hypothesized to be encoded in phonology as markedness. However, this explanation then would not apply to the palatalization of voiced velar stops. Across languages, if a language palatalizes [g] before front vowels, it also palatalizes [k] (Moreton & Pater, 2012b). However, the participants in Wilson (2006)'s experiment did not generalize a palatalization pattern from voiced velar stops to voiceless velar stops. He attributed this to another perceptual factor, namely the perception of [voice] in stops and suggested that the interaction of two perceptual factors in one phonological pattern make it difficult to establish a direct relation between the learning of this pattern and its typology. There is a plausible alternative explanation for Wilson (2006)'s findings concerning palatalization of voiceless velars, furnished by Kapatsinski (2013a), which is based on phonetic similarity rather than on effects of ease of articulation and ease of perception. Kapatsinski (2013a) assumed that listeners overgeneralize their learned pattern based on phonetic similarity: As [e] is phonetically between [i] and [a], participants trained on the palatalization in the context of [e] and on the non-palatalization in the context of [a] will also admit the palatalization in the context of [i] due the phonetic similarity between [e] and [i]. I will describe this explanation in detail below in chapter 1.1.2.

In short, even though some findings support a connection between ease of perception and ease of articulation of a phonological pattern, its learnability and typology, others do not. This raises the question whether there is a connection between the ease of perception or the ease of articulation of a pattern, its learnability on the one hand, and its typology on the other hand. I will investigate exactly this question and I will suggest a further phonetic factor affecting phonological learning:

If Kapatsinski (2013a)'s reanalysis of Wilson (2006)'s results of palatalization of voiceless velar stops is on the right track, generalizations of a phonological pattern may be governed by phonetic similarity.

In this dissertation I will consider the effects of ease of articulation and ease of perception on phonological learning and phonological generalizations separately. I will do this by investigating the learnability of a vowel nasalization pattern for vowels of different height (see chapter 2.4) as ease of articulation and ease of perception make different prediction regarding a learning advantage and a generalization advantage for vowel nasalization of vowels of different height. Whereas the nasalization of low vowels is the articulatory easiest, the nasalization of high vowels is the perceptually easiest. As the influence of these biases may show up in perception as well as in production, I will conduct not only a perceptual task but also a productive task (see chapter 3). Doing so, I can differentiate the contribution of ease of perception and ease of articulation in the perception of the learners as well as in the articulation of the learners.

1.1.2 Phonetic similarity

Another aspect of phonetics that may affect the learning of phonological patterns is their phonetic similarity (Kapatsinski, 2013a). As noted above in chapter 1.1 phonetic similarity refers to the similarity between certain sounds or patterns. They are phonetically similar to each other when they have similar phonetic features or when they are acoustically and perceptually close to each other. There are two subtypes: Phonetic similarity will be discussed in this chapter, whereas the perceptual similarity to native language phonemes will be discussed separately in chapter 1.1.3 below.

Even though the literature focused on the role of phonological similarity – not phonetic similarity – in learning alternations, there is some work on the role of phonetic similarity in learning phonotactic generalizations. Knowledge of phonotactics of adults is best captured by means of natural classes of phonological features (Daland, Hayes, Garellek, White, Davis & Norrmann, 2011; van de Vijver & Baer-Henney, 2012). Natural classes can be thought of as phonologized phonetic similarities (Clements, 1985, 2009; Clements & Hume, 1995; Cristià et al., 2013; McCarthy, 1988; Uffmann, 2011). Thus, phonetic details that are used to group segments in natural classes are incorporated in phonology as features. Similarity is an important factor in generalizations concerning existing morphophonological patterns in a language (Albright & Hayes, 2003; Blevins & Blevins, 2009; Ernestus & Baayen, 2003). Albright & Hayes (2003) found that native speakers of English inflect pseudowords (verbs) based on the similarity of the pseudowords to existing verbs. If a pseudoword resembles existing verbs with irregular past tenses, it is likely to be inflected irregularly itself. The phonetic similarity is calculated based on the natural class based theory by Broe (1993). Ernestus & Baayen (2003) similarly found that Dutch native speakers inflect pseudowords based on their phonetic and phonological similarity to existing verbs in the mental lexicon. The explanation of the findings of Albright & Hayes (2003) and Ernestus & Baayen (2003) can be paraphrased as follows: Learners take a pattern they know as a particular good instance of a morphophonological form, and a novel pattern will be evaluated based on its resemblance to known forms. The more a novel pattern resembles a known one, the more likely it will be given a morphophonological form similar to the known one.

Daland et al. (2011) investigated the knowledge of English native speakers of the well-formedness of English words, like *brick*, and pseudowords, such as *blick*, and non-words, such as *bnick*. They created pseudowords and non-words and asked participants to assess their well-formedness. Based on an analysis of the onsets in a corpus of English words as sequences of natural classes, they found that the best explanation of the assessments was the frequency of the sequence of natural classes. A similar role for natural classes, and thus for phonetic similarity, has been found in the phonotactics of onset clusters in English (Hayes & Wilson, 2008) and in German (van de Vijver & Baer-Henney, 2012). van de Vijver & Baer-Henney (2012) conducted a rating study in which German native speakers were asked to judge the well-formedness of onsets in pseudowords. It turned out that German native speakers rated pseudowords with onsets with sequences of natural classes that are frequent in German best: The rarer the sequence of natural classes in German, the worse the rating. In addition to that, onsets attested in the German lexicon were rated higher than pseudowords with unattested onsets. Among the unattested onsets unmarked onsets – those with a rising sonority profile – were given a higher rating than marked onsets – those with a sonority plateau or a falling sonority. A phonotactic learner model confirmed these results and further showed that distinctive features are used to derive the preferred onsets from the lexicon.

There is also evidence that shows that phonetically similar patterns are preferred in learning over phonetically less similar patterns, even though this difference in preference is not reflected typologically. For example, Cristià & Seidl (2008) investigated whether infants use phonetic details to group segments in natural classes, and whether they can use these groupings to learn about the phonotactics of a miniature language. Cristià & Seidl (2008) taught one group of infants that a word always starts with a plosive or a nasal – the natural class of sounds that are [$-$cont]. Sounds that are [$-$cont] are pronounced with a complete closure in the mouth. Another group was taught that a word always starts with a nasal or a fricative. Nasals and fricatives cannot be united in a natural class by means of one or more phonological features. They found that it is more difficult to learn words that start with segments drawn from a phonetically arbitrary set of sounds. Even though a preference for words starting with a plosive or a nasal as opposed to words starting with a nasal or a fricative can be explained phonetically, there is no evidence that this pattern is reflected typologically (Moreton & Pater, 2012b).

Cristià et al. (2013) investigated different measures of similarity in learning phonotactic patterns. French native speakers were exposed to words from an artificial language. The segments in the onsets of the words were classified along different similarity dimensions. For example, onsets with [g] are compared to other onsets that are similar along different dimensions: The onset [b] belongs to the narrowest natural class, differing only in place of articulation, the onset [k] belongs to a near natural class, differing only in voice, and the onset [p] belongs to a far natural class, differing in voice and place. During the exposure phase participants were asked to judge the well-formedness of these words. In a subsequent test phase they heard novel words, and were additionally asked to judge their relative frequency, even though they had heard all words the same number of times. The words that were judged as most frequent were those from the narrowest and the near natural class of the words presented during exposure. Cristià et al. (2013) concluded that generalizations were made on the basis of similarity between the onsets of the words in exposure and the onsets of the words in the test phase.

The above-mentioned studies investigated the role of phonetic similarity in learning and generalizing phonotactics. There is, however, less work on the role of phonetic similarity in learning and generalizing (morpho)phonological alternations or patterns. Examples in which phonetic similarity is used as an explanation for learning a phonological alternation and for generalizations of a learned phonological alternation are the study of Skoruppa et al. (2011) as well as the above-mentioned alternative explanation for Wilson (2006)'s findings concerning palatalization of voiceless velars.

Skoruppa et al. (2011) exposed participants to phonological alternations in which either one, e.g. place of articulation, two, e.g. place of articulation and manner of articulation, or three features, e.g. place of articulation, manner of articulation and voicing, changed. Participants were better in learning a one-feature-change than a two- or three-feature-change. Skoruppa et al. (2011) infer that the phonetic distance based on features influences phonological learning: The larger the phonetic distance between two alternating sounds is, the more difficult this alternation is to learn. Note that in this case phonetic similarity is confounded with structural simplicity so that these results may be due to a simplicity bias.

Whereas Wilson (2006) interpreted his results in the light of ease of articulation and ease of perception, Kapatsinski (2013a) interpreted them based on phonetic similarity. Wilson (2006) found, to briefly recapitulate, that learners, who were taught that [k] is pronounced as [t͡ʃ] before [e] and as [k] before [a], generalized that [k] is pronounced as [t͡ʃ] before [i] and that learners, who were taught that [k] is pronounced as [t͡ʃ] before [i] and as [k] before [a], tended not to generalize that [k] is pronounced as [t͡ʃ] before [e]. Kapatsinski (2013a) pointed out that [e] is perceptually between [i] and [a], and he argued that listeners always generalize slightly beyond observable patterns. This means that after being taught that [ke] palatalizes to [t͡ʃe] and that [ka] remains [ka], participants have evidence for [t͡ʃe] (provided to the participants) and [t͡ʃi] (generalized by the participants based on perceptual similarity between [e] and [i]). The learners in this group, therefore, have evidence for [t͡ʃe], [t͡ʃi] and [ka], and explicit evidence against [ke]. Learners who have learned that [ki] palatalizes to [t͡ʃi] have explicit evidence for [t͡ʃi] (provided to the participants) and [t͡ʃe] (generalized by the participants on the basis of perceptual similarity between [e] and [i]). They also have explicit evidence for [ka], and implicit evidence for [ke] (because [e] and [a] are perceptually similar). Learners in this group, therefore, have evidence for [t͡ʃi], [ka] and for [t͡ʃe] and [ke], and are consequently uncertain as to the correct consonant preceding a mid vowel. This is due to mid vowels' phonetic similarity to high vowels and to low vowels, which both give evidence for different consonants. Thus, it is possible that the velar palatalization results do not support the claim that the ease of articulation or the ease of perception of a phonological pattern affords it a learning advantage. Instead, phonetic similarity might fully explain the results.

Further evidence for a preference of phonetically similar patterns comes from Steriade (2001b, 2009)'s P-map theory (perceptibility map). She argued that speakers have detailed knowledge of the perceptual similarity between sound patterns and that learners prefer phonological alternations with small perceptual changes. White (2013) investigated this experimentally with artificial language learning experiments by studying the learnability of a saltatory alternation – an alternation that involves a two-feature change, whereas intermediate patterns with a one-feature change remain unaffected from this alternation – with infants and adults. An example for a saltatory alternation is the change from voiceless plosive [p] to voiced fricative [v], for which the voiceless fricative [f] and/or the voiced plosive [b] have to be skipped, which means that [f] and [b] would be unaffected from this alternation and would never change to [v]. In one of his experiments, White (2013) trained one group of participants on a saltatory alternation ([p]–[v] and [t]–[ð]) without exposing participants to the intermediate sounds [b], [f], [d] and [θ]. Another group – the control group – was trained on a non-saltatory alternation ([b]–[v] and [d]–[ð]) and was not exposed to [p], [f], [t] and [θ]. If participants in the saltatory condition stick to the saltatory alternation, they will not change the untrained intermediate sounds. However, if they have a bias to disprefer saltatory alternations – or in other words, a bias to prefer alternations with small perceptual changes –, they will change the untrained intermediate sounds. He found that learners are indeed biased to prefer alternations between similar sounds, which means that participants in the saltatory condition also changed the intermediate sounds – so that the change was smaller – although there was no direct evidence for this change during training. The infants' reaction also indicate that they generalized a learned rule between dissimilar sounds to a pattern with similar sounds, but not vice versa. This experiment thus showed that participants who learned an alternation between dissimilar sounds generalized it to an alternation between more similar sounds, which is in line with Steriade (2001b, 2009)'s P-map theory which says that learners prefer small perceptual changes.

In sum, phonetic similarity is used by adults as well as by infants to learn about phonological patterns. They make use of phonetic similarity as instantiated in natural classes when they generalize from observed to novel forms. For generalizations from observed to novel forms in a morphophonological artificial language learning experiment this would mean that generalizations to novel forms are made on the basis of the phonetic similarity of these novel forms to known forms from training. This phonetic similarity is a further phonetic factor affecting phonological learning, adding to the often discussed and criticized phonetic factors ease of perception and ease of articulation as proposed within the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006). Inspired by Kapatsinski (2013a)'s reinterpretation of Wilson (2006)'s findings I propose the following hypothesis: Learners will take a pattern they know (from training) as a particular good example of a phonological category, and novel patterns will be evaluated on the basis of their phonetic resemblance to known patterns. The more a novel pattern resembles a known pattern, the better they will rate the novel pattern. This means that the phonetic similarity between phonological patterns affects their generalizability.

In this dissertation I will consider the effect of phonetic similarity on phonological generalizations (see chapters 2.4 and 3). This phonetic similarity is not necessarily connected with typology. I propose that learners make phonetic generalizations in which phonetically similar patterns are learned easier than dissimilar patterns. As phonetic similarity does not necessarily dovetail with ease of articulation or ease of perception, I will be able to differentiate the contribution of phonetic similarity from that of ease of perception and ease of articulation.

1.1.3 Perceptual similarity to native phonemes

The above-mentioned phonetic similarity and its effect on phonological generalizations always assumes that the learning process is taking place as expected. Participants listen to certain stimuli, perceive and recognize the pattern in the way the experimenter expects, and generalize it to other stimuli that are again perceived in the way the experimenter wants them to be perceived. However, it might be the case that unfamiliar sounds or pattern which are under investigation in those experimental studies are perceived differently – namely not as the unfamiliar sound or pattern but as a similar sounding native sound or pattern. Although the perceptual similarity to native phonemes is based on acoustics, and in some part also on articulation, I will call it *perceptual similarity* as the parameters to establish that sounds are similar to native phonemes or not are measured based on results of a perceptual task. This, however, does not mean that the sounds under investigation are only similar in terms of perception.

Theories about perception and learning of non-native sounds and contrasts were developed by Kuhl (1991) (native language magnet), Best (1995) (perceptual assimilation model) and Flege (1995) (speech learning model). They all agree that the ability to perceive non-native sounds as native or non-native depends on the perceptual distance between non-native and native sounds. If native and non-native sounds are perceptually similar, the non-native sound will be perceived as an example of a native sound. If both sounds are perceptually different, the non-native sound will be perceived as such. Concerning the perception of non-native sounds, Best (1995) differentiates between three types of assimilation patterns: Non-native sounds can be assimilated to a native category, which means that they are either perceived as an ideal, good or poor exemplar of that category, or they can be assimilated as an uncategorizable speech sound, which means that they are perceived as non-native speech sounds, or they can be unassimilated to speech, which means that they are perceived as non-speech sounds. She further differentiates several types of assimilation patterns for non-native contrasts and their discrimination ability, e.g. if both non-native sounds are perceived as another native category (termed two-category assimilation), discrimination will be very good. The same is true if one non-native sound is perceived as native, the other one as non-native sound (termed uncategorized versus categorized). If both sounds are perceived as the same native category, the discrimination ability will depend on the proximity to the native sound: If this proximity differs for both sounds (termed category goodness difference), discrimination will be better than if both sounds are equally distant from the native sound (termed single category assimilation). These perceptual differences allow predictions regarding learnability.

Kuhl & Iverson (1995) and Flege (1995) assume that a perceptually more distant non-native sound can be learned easier than one that is perceptually close to a native sound – not only this particular non-native sound but also the contrast between this sound and native sounds. However, Aoyama, Flege, Guion, Akahane-Yamada & Yamada (2004) consider another possibility into account and additionally assume a learning advantage for a non-native sound which is very similar to a native sound. In this case, the non-native sound is directly mapped onto an already existing native sound. No new sound category has to be established. However, we have to keep in mind that in this case no new non-native sound is "learned" but it is misperceived and "mislearned" as a native sound.

These theories were extensively tested with various cases. The classic example is the perception of the English /l/-/r/-contrast by Japanese native speakers. In American English there are two approximants [l] and [ɹ] (Ladefoged, 2001), but in Japanese there is just the tap [ɾ] (Vance, 1987). Several studies showed that Japanese listeners assimilate both English sounds to the Japanese sound and suggested a 2-to-1 mapping as there is only one perceptually similar sound in Japanese (Best & Strange, 1992; Guion, Flege, Akahane-Yamada & Pruitt, 2000; Komaki, Akahane-Yamada & Choi, 1999; Takagi, 1993). However, the English consonants differ in terms of their goodness of fit with respect to the Japanese sound: The studies of Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Kettermann & Siebert (2003), Komaki et al. (1999) and Takagi (1993) showed that English [l] is considered to be more similar to Japanese [ɾ] than English [ɹ]. This means that the perception and the learning of English [ɹ] should be easier for Japanese native speakers than that of English [l]. As English [ɹ] is perceived as less similar to the native category than English [l] is, it should be easier for learners to identify it as a non-native sound, and thus to establish English [ɹ] as a new category. This is what was indeed found: Bradlow, Pisoni, Akahane-Yamada & Tohkura (1997), Flege, Takagi & Mann (1996) and Sheldon & Strange (1982) found better identifications for English [ɹ] over [l] by Japanese native speakers and Aoyama et al. (2004) found a learning advantage for [ɹ] over [l]. Aoyama et al. (2004) studied the perception and production of the /r/-/l/-contrast with Japanese adults and children at two time points (with one year in between) and found perceptual learning and an improvement in production for [ɹ] for children. Adults' scores in both production and perception, however, did not improve significantly from time point 1 to time point 2 – probably because the Japanese adults had an initial advantage due to English lessons in school.

Another example is the study of Polka (1995). She tested the perception of the German vowel contrasts /u/-/y/ and /ʊ/-/ʏ/ with native speakers of English. English has only back round vowels (/u/ and /ʊ/) and no front round vowels (/y/ and /ʏ/). The performance of the English listeners was better for the tense contrast /u/-/y/ than for the lax contrast /ʏ/-/ʊ/. In a second task the participants were asked to identify the German vowels as English ones and to rate them according to their goodness of fit. The English listeners perceived German /u/ and German /y/ both as English /u/, whereas German /u/ was considered a better instance of English /u/ than German /y/. This difference in similarity to a native sound explains the good performance in the perception of the tense contrast: English listeners identified the German tense contrast as good and poor instances of an English vowel, and performed thus very well. The results for the lax vowels were similar, but not as clear as for the tense vowels, with German /ʊ/ being considered perceptually more similar to English /u/ than German /ʏ/. This again explains the poorer performance of the perception of the lax contrast: English listeners were not able to clearly map the German contrast onto good and poor examples of one English vowel – some of the contrast pairs were equally well (or poor) so that their performance decreased. Taken together the studies of Aoyama et al. (2004) and Polka (1995) showed that the perception of non-native sounds and contrasts depends on the perceptual similarity to native sounds and that the difficulty in the perception of contrasts is influenced by the way in which non-native sounds are mapped onto native ones.

In this dissertation I will explore the role of perceptual similarity in phonological learning. I will focus on the perceptual similarity between alternating sounds (as proposed by phonetic similarity (see chapter 1.1.2)) and on the perceptual similarity between native and non-native sounds – testing a possible interference of L1 with the latter. As the interference of L1 may depend on how reliable the evidence for a morphophonological alternation is in the learning situation, I will conduct not only an artificial language learning experiment with evidence for a morphophonological alternation (see chapter 4.4) but also without evidence for a morphophonological alternation during training (see chapter 5). Doing so, I will on the one hand be able to differentiate the contribution of language-general perception and language-specific perception in phonological learning and on the other hand I will be able to demonstrate whether learners can transfer one feature from a trained pattern to novel items without knowing about the relation between phonological forms or whether the relation between phonological forms has to be known to generalize a trained pattern to novel items.

1.2 The use of artificial language learning experiments

As shown with the experimental studies mentioned above, the role of learning biases is usually tested by means of artificial language learning experiments. Thus, I will make use of this method in the experiments designed for my dissertation as well. An artificial language is a constructed miniature language which does not exist in reality but which may be a possible language (Culbertson, 2012). This miniature language includes a phonological pattern that is not present in the participants' native language (Glewwe, 2019). Consequently, differences in learning are assumed to be due to the bias under investigation (Glewwe, 2019). Note that artificial language learning experiment are not only used to test phonological learning but also to test e.g. morphological, semantic or syntactic learning (Merkx, Rastle & Davis, 2011; Mintz, 2002; Morgan & Newport, 1981).

Artificial language learning experiments in phonology usually consist of a learning or training phase in which participants are exposed to auditory stimuli which contain a certain alternation or pattern. These auditory stimuli can be supported visually by images or symbols to convey a meaning, e.g. express plural by means of two symbols as in Peperkamp et al. (2006) and White (2013). There are, however, also experiments without visual support as e.g. in Finley (2009). Participants are most often not expected to learn the semantic meaning of the form or think about what they hear (Culbertson, 2012). They are rather asked to just listen and look at the screen as in Baer-Henney & van de Vijver (2012) and Finley (2009) – except for e.g. in the study of Pycha et al. (2003) in which participants were asked to find out how the plural is formed in the artificial language constructed for this study. To make the experiment more exciting some researchers also give the experiment a gaming character as in Wilson (2006). In most of the training phases no negative input is given to the participants, which means that participants are only exposed to correct forms so that they do not learn what is wrong. This method mimics the implicit learning of a natural learning situation of a child that mainly hears what adult speakers are saying in everyday life (Glewwe, 2019; van de Vijver & Baer-Henney, 2014). In the artificial language learning experiments conducted for this dissertation the auditory stimuli will be supported visually and the participants will not be asked to figure out a rule.

After the training phase the test phase follows. Sometimes there is an intermediate phase between training and test which can vary from a two minute break to work on pencil-and-paper math problems (Wilson, 2006) to twelve hours of sleep (Earle, Landi & Myers, 2017; Martin & Peperkamp, 2020), but in most experiments the test directly follows training. The test can include various tasks. It can, for example, be a productive task as in Baer-Henney, Kügler & van de Vijver (2015b), Peperkamp et al. (2006) and Wilson (2006), a perceptual forced-choice task as in Finley (2009, 2012), Moreton (2008) and White (2013) or a written task as in Becker et al. (2012) to name just a few. This test phase aims at investigating whether the pattern the participants had been exposed to was learned and – depending on the structure and aim of the experiment – how well participants generalized the learned pattern to other patterns. I will make use of a perceptual forced-choice task and a productive task which directly follows training in the artificial language learning experiments presented in this dissertation to investigate whether the effects of biases which are grounded in perception and of biases which are grounded in articulation show up in the learners' perception and in the learners' articulations.

One way of testing this is by using an artificial language in which learners are taught a phonological pattern that implies the presence of another pattern and investigate whether the implication is learned (Finley, 2012). Another way of testing this is by using an artificial language with two or more patterns which differ in their ease of articulation, ease of perception or phonetic similarity (or any other aspect of interest) and investigate whether the easier or more similar pattern is learned better than the more difficult or dissimilar pattern (Skoruppa & Peperkamp, 2011). More insights in the process of learning gives the way learners generalize the learned pattern to other patterns. To test this, Wilson (2006) used the poverty-of-the-stimulus method: Participants are trained on only one pattern but they are tested on more than one pattern. In most experiments participants are divided into learning groups which are all trained on a different pattern, e.g. on the velar palatalization before [i], [e] or [a]. In the subsequent test phase participants of all groups are then tested on all patterns, e.g. on the velar palatalization before [i], [e] and [a], in order to test how learners generalize their learned rule from training to novel patterns at test. As this method was used successfully in several artificial language learning experiments (Baer-Henney, 2015; Baer-Henney et al., 2015a; Cristià & Seidl, 2008; Finley, 2009, 2012; White, 2013, 2014; White & Sundara, 2014; Wilson, 2003, 2006; Zuraw, 2007), I will make use of this method as well.

The reasons why artificial language learning experiments are used to explore the role of learning biases and other effects on learning phonological patterns are diverse. By means of artificial languages one can create a simple linguistic context – a miniature language – in which one can compare one or more patterns directly without any interference. One can control for confounding factors that might influence the learning process so that one can investigate only the particular pattern or alternation under investigation (Culbertson, 2012). This is not possible in a natural learning environment in which the linguistic context is very complex – with several phonological processes taking place and affecting one sound at the same time. In a natural learning environment there are also many confounding factors, such as relative frequency of the occurrence of one pattern in the native language, the different degree of familiarity with a pattern by different participants with different language skills or different phonetic contexts, which one cannot control for. Another advantage of the use of artificial language learning experiments is that one can investigate the complex process of learning in a relatively short time at one place without much effort. Testing the same process of learning in natural language acquisition would be of more effort and would take much more time, e.g. longitudinal study.

However, just as any method, the artificial language learning paradigm also has drawbacks. Critics argue that in artificial language learning experiments participants may use different learning strategies from those used in natural language acquisition (Culbertson, 2012). This is a problem as the experiments aim at making statements about the learning process in general. Furthermore, it is not clear whether a very short training phase of only five to ten minutes in a laboratory setting has the same effect as a longer learning process of several weeks, months or years (Glewwe, 2019). Consequently, according to Glewwe (2019) one has to keep in mind that the presence of a bias in an artificial language learning experiment does not imply the presence of this bias in natural language acquisition and that the absence of a bias in an artificial language learning experiment does not imply the absence of this bias in natural language acquisition. Glewwe (2019) therefore suggests conducting artificial language learning experiments and natural language studies to compare the results. For a detailed review on artificial language learning experiments investigating different types of learning biases see Culbertson (2012) and Moreton & Pater (2012a,b).

Despite the drawbacks mentioned above, I think that artificial language learning experiments are most suitable to tease apart different influences on the learning and generalization process. Without artificial languages one will not be able to investigate whether ease of perception, ease of articulation or phonetic similarity influences learning a particular pattern and how learners generalize their learned rule to another pattern.

1.3 Justification for the use of vowel nasalization in my experiments

To test whether ease of perception, ease of articulation or phonetic similarity influences learning a particular pattern and how learners generalize the learned pattern to another pattern, a phonological pattern is needed that makes different predictions based on these three factors. Therefore, I will test the learnability of vowel nasalization of three different heights – low [a], lower mid [ɛ] and high [i] – and how the learned pattern is generalized to the other vowel heights. Vowel nasalization is a process by which the velum is lowered during the articulation of a vowel. Due to this lowering of the velum the nasal cavity is connected to the oral cavity, which means that both cavities resonate together (Ohala, 1975) – in contrast to the articulation of oral vowels which is characterized only by resonances of the oral cavity. This connection of oral and nasal cavity has acoustic consequences, e.g. raised frequency of the first formant (F1) or reduced amplitude of F1, and likewise perceptual consequences (Ohala, 1975). More information on the acoustics and the perception of vowel nasalization are provided below and in chapter 2.2 and 2.3.

All participants were German native speakers and because German native speakers are not familiar with phonemic vowel nasalization (Laeufer, 2010; Wiese, 1996) a possible influence of frequency distributions of nasal vowels in their L1 can be excluded. There are, however, French loan words in German in which nasal vowels are imitated by German native speakers (Laeufer, 2010). Thus, the test case of vowel nasalization is not only suitable to investigate how small phonetic details influence the perception of the learners but also the production of the learners. To do this, I will test whether the exposure to vowel nasalization leads to better imitations of nasal vowels in French loan words by German native speakers.

I will now describe how the nasalization differs for different heights. In terms of vowel nasalization the articulation of low nasal vowels is easier than that of non-low nasal vowels. The reason for the ease of articulation of low nasal vowels lies in the anatomic connection between the muscles used for lowering the velum and the muscles used for lowering the tongue body. The palatoglossus muscle connects the tongue with the velum. A contraction of the palatoglossus causes a lowering of the velum, which means that the velopharyngeal port is open and a nasal sound is articulated (Bell-Berti, 1993). The lowering of the tongue body for the articulation of low vowels is achieved by the hyoglossus muscle. Both muscles, palatoglossus and hyoglossus, are anatomically connected with each other, which means that a lowering of the tongue body by the hyoglossus to produce a low vowel pulls the palatoglossus down which causes the lowering of the velum and thus the nasalization of the low vowel (Ohala, 1975). This is the reason why low vowels are easier to nasalize than high vowels for whose production additional muscles would be required.

Let us now turn to perception. The distinction between high nasal and oral vowels is easier to perceive than the distinction between non-high nasal and oral vowels. The reason for the greater perceptual difference between oral and nasal high vowels compared to oral and nasal non-high vowels lies in the different degree of acoustic modification due to the nasalization. The most prominent acoustic consequences of nasalization are the reduction of the first formant's amplitude (Delattre, 1954; Delvaux, 2009; Fant, 1960; House & Stevens, 1956; Macmillan, Kingston, Thorburn, Walsh Dickley & Bartels, 1999; Pruthi & Epsy-Wilson, 2004; Schwartz, 1968; Stevens, 1998) as well the introduction of additional resonances and anti-resonances in the vicinity of the first formant due to the nasal cavity as an additional resonator (Chen, 1995, 1997; House & Stevens, 1956; Kingston & Macmillan, 1995; Mermelstein, 1977; Schwartz, 1968; Stevens, 1998). Measuring nasality is thus possible by investigating changes in amplitude in the spectrum, e.g. measuring spectral tilt, which is the decrease of amplitude in neighboring frequencies. This spectral tilt is very similar in oral and nasal non-high vowels, but less so in oral and nasal high vowels (Schwartz, 1968; Styler, 2015). Consequently, non-high oral and nasal vowels are acoustically more similar than high oral and nasal vowels. Previous acoustic studies confirmed that low oral and nasal vowels are acoustically more similar than non-low oral and nasal vowels (Ohala, 1975; Schwartz, 1968). This acoustic difference is reflected in perception. Native speakers of American English identified high nasal vowels best, followed by mid and low nasal vowels (Bond, 1976). These results are supported by a study of House & Stevens (1956) who showed that the lower the vowel, the more velum lowering is required to judge a vowel as nasal by native speakers of American English even though they are familiar with nasal vowels (Hayes, 2009).

Another reason to choose vowel nasalization is the phonetic similarity among vowels of different heights. Phonetic similarity refers to how similar certain sounds or patterns are. Sounds or patterns are phonetically similar when they have the same phonetic features or when they are acoustically close to each other. In the case of vowel nasalization this phonetic similarity is independent of nasalization. Low [a] and mid [ɛ] are phonetically more similar to each other than mid [ɛ] and high [i]. The reason for this is that [ɛ] is a lower mid vowel, with formant frequencies closer to low [a] than to high [i]. The same is true for the nasal vowels: [ɛ̃] is phonetically more similar to [ã] than to [ĩ]. To test whether this is also true for the stimuli I used in the experiments on vowel nasalization I did a short acoustic analysis (see chapter 2.1). The results confirm my expectations: (Lower) Mid [ɛ] is phonetically more similar to low [a] than to high [i] – independent of whether the vowels are oral or nasal.

A further reason to investigate vowel nasalization is perceptual frequency. It refers to how often we are exposed to a phonetic detail and consequently how much memory traces there are for a given phonetic detail (Goldinger, 1998; Zellou, Dahan & Embick, 2017). As low-frequency words are represented by fewer traces, the influence of a low-frequency prime on the activation of a single representation is greater – due to missing concurrent words – than the one of a high-frequency prime, and this induces a better imitation of this prime in production (for details see chapter 3.1). In the case of vowel nasalization high nasal vowels correspond to low-frequency primes and mid and low nasal vowels correspond to high-frequency primes for German native speakers. This is because German native speakers perceive (strongly) nasalized vowels most often in French loan words, and as there are no high nasal vowels in French (Fagyal, Kibbee & Jenkins, 2006), German native speakers are hardly exposed to high vowel nasalization.

A fifth factor is often said to influence phonological learning as well: structural simplicity (Moreton, 2008, 2012). Something is structurally simple when it involves only few distinctive/phonetic features. This factor, however, is not under investigation here. At first sight, structural simplicity might correlate with phonetic similarity that is partly based on phonetic features, but the two factors focus on different aspects of similarity. Whereas structural simplicity refers to the number of features of the tested sound patterns, phonetic similarity refers to the similarity between features while keeping the number of features constant. Let me illustrate this with the example of vowel nasalization.

I will investigate the learnability and the imitation of vowel nasalization of vowels of different heights. These vowels do not differ in their structural simplicity in terms of distinctive features (Pulleyblank, 2011; Uffmann, 2011): When a vowel is nasalized, the feature [nasal] is added to a former (default) oral vowel. Adding the feature [nasal] to a low vowel is of the same structural complexity as adding it to a mid vowel or to a high vowel as illustrated in Figure 5. Thus, studying the learnability of vowel nasalization, its generalizations and imitations avoids the confounding factor structural simplicity.

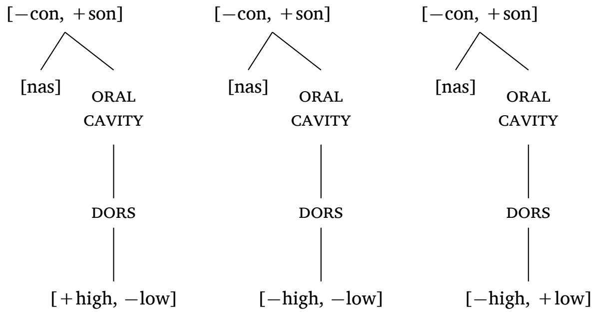

Figure 5: Structure of high, mid and low nasal vowels.

To sum up, vowel nasalization is used as an experimental test case to investigate whether phonological learning is influenced more by ease of articulation or by ease of perception, whether learners base their generalizations more on ease of articulation, ease of perception or phonetic similarity and whether learners base their imitations more on ease of articulation, ease of perception, phonetic similarity or perceptual frequency. Ease of articulation predicts a learning advantage for low nasal vowels as well as better imitations of and more generalizations to low vowels for all learning groups, ease of perception predicts a learning advantage for high nasal vowels as well as better imitations of and more generalizations to high vowels for all learning groups, phonetic similarity predicts different imitations and generalizations depending on the learning group (e.g. better imitations of and more generalizations from lower mid [ɛ] to low [a] than to high [i], but better imitations and more generalizations from [i] to lower mid [ɛ] than to low [a]) and perceptual frequency predicts better imitations of high vowel nasalization than of mid or low vowel nasalization. For detailed predictions see chapters 2.4 and 3.1.

1.4 Justification for the use of retroflexion in my experiments

To test whether ease of perception or phonetic similarity between native and non-native sounds influences learning a particular pattern and how learners generalize the learned pattern to another one, a phonological pattern is needed that makes different predictions based on these two factors. Therefore, I will test the learnability of retroflexion of three different consonant manners – plosive [t], fricative [s] and lateral approximant [l] – and how the learned pattern will be generalized to the other consonant manners.

Retroflex describes a place of articulation, such as bilabial, dental or velar based on the IPA chart. In this case *retroflex* refers to a place of articulation between the alveolar ridge and the palatum, which is rather variable in contrast to other places of articulation that have a fixed place. Additionally, there is more into *retroflex* than only the articulation in the palatoalveolar region. Retroflex sounds are traditionally described as sounds articulated with a backward bending of the tongue (Trask, 1996). However, in order to reflect the complexity of retroflex sounds, Hamann (2003b) proposed a new definition of retroflex that includes four articulatory characteristics – apicality, posteriority, sublingual cavity and retraction – and suggested *retroflex* to describe a gesture and not a fixed place of articulation. These articulatory characteristics have acoustic consequences that can be observed in the formant frequencies and transitions as well in the spectral shapes of retroflex consonants. Retroflex consonants most likely have a lower F3 than their alveolar counterparts (Ladefoged & Maddieson, 1996; Narayanan & Kaun, 1999) and formant transitions from the retroflex consonants into a vowel often show a decreasing F2 and an increasing F3 (Stevens, 1998; Stevens, Keyser & Kawasaki, 1986). Generally, the formant transitions from retroflex consonants are quite short compared to those from alveolar consonants (Stevens et al., 1986). The spectral shape of a retroflex consonant shows an abrupt decrease of amplitude in the higher frequencies as well as concentrated energy in the lower frequencies (Dart & Nihalani, 1999). The frication noise most often starts lower in retroflex fricatives than in alveolar fricatives (Ladefoged & Maddieson, 1996; Stevens & Blumstein, 1975). Moreover, there is a strong mid-frequency peak in retroflex consonants that is lower than in alveolar consonants (Ladefoged & Maddieson, 1996). For a detailed overview of the phonetics and phonology of retroflex consonants see Hamann (2003b) and for information on the perceptual consequences of retroflexion see chapter 4.2.

To investigate the different effects of the phonetics of retroflexes on their learning, I will test German native speakers because German native speakers are not familiar with retroflex consonants so that a possible influence of frequency distributions of retroflex consonants in their L1 can be excluded. German native speakers are, however, familiar with the postalveolar fricative [ʃ] that is perceptually and acoustically similar to the retroflex fricative [ʂ]. There are no such similar sounds for the other retroflex consonants in German (Wiese, 1996). This enables me to test a possible interference of one phonetically similar L1-sound on learning and generalizations.

I will now describe how the retroflexion differs for different manners. Ease of perception refers to how easy it is to perceive a certain sound or pattern. The perception of retroflex fricatives is easier than the perception of retroflex plosives or nasals (Stausland Johnsen, 2012). Stausland Johnsen (2012) conducted a perception experiment with native speakers of Norwegian. Based on the perceptual data he gained he calculated the perceived distances for the alveolar-retroflex consonant pairs. Stausland Johnsen (2012) found that the perceived distance between alveolar and retroflex fricatives is greater than that of alveolar and retroflex plosives or nasals. He explained his data with similar data from German regressive place assimilation by Kohler (1990). Kohler (1990) found that alveolar plosives and alveolar nasals totally assimilate to a following labial consonant, whereas alveolar fricatives do not. These findings were explained with the acoustic weakness of nasals and plosives after vowels. According to Kohler (1990) they are relatively less salient so that omitting them would not be that remarkable as omitting fricatives that are acoustically more salient than plosives or nasals. Thus, the high acoustic salience of fricatives compared to other consonant manners like plosives or nasals is the reason why modifications of fricatives lead to a perceptually greater change than those of other consonant manners. Thus, the perceptual distance between fricative pairs, e.g. [s]-[ʂ], is greater and consequently perceived better than that between other consonant manner pairs, e.g. [t]-[ʈ].

Phonetic similarity refers to how similar certain sounds or patterns are. Sounds or patterns are phonetically similar when they have the same phonetic features or when they are acoustically close to each other. In the case of retroflexion the phonetic similarity refers to the similarity between retroflex consonants and German native sounds. If native and non-native sounds are perceptually similar, the non-native sound will be perceived as an example of a native sound. If both sounds are perceptually different, the non-native sound will be perceived as such. My intuition as a German native speaker and the results of a perception experiment (see chapter 4.3) agree that the German postalveolar fricative [ʃ] is perceptually very close to the retroflex fricative [ʂ] for German native speakers. There are no perceptually similar German sounds for the other retroflex consonant manners under investigation. Thus, the retroflex fricative is perceptually more similar to a native category than the other retroflex consonants (for details see chapter 4.3).

Again retroflex consonants of different manners do not differ in their structural simplicity in terms of distinctive features. When an alveolar consonant is retroflexed, the feature [anterior] underneath the CORONAL node changes from [+anterior] to [−anterior] (Chomsky & Halle, 1968; Clements, 1991). All other features are unaffected by retroflexion. Thus, changing the value of the feature [anterior] from [+anterior] to

[−anterior] in an alveolar fricative is of the same structural complexity than changing the value of the feature [anterior] from [+anterior] to [−anterior] in an alveolar lateral approximant or in an alveolar plosive as illustrated in Figure 6. Some researchers argue for additional features such as [+back] (Gnanadesikan, 1993) or [+high] (Pulleyblank, 1989) to discriminate retroflex consonants from non-retroflex consonants. However, as this requirement is language-specific and as the addition or change of these feature values does not differ across consonant manners, I avoided this in the feature representation in Figure 6 for reasons of simplicity. Thus, studying the learnability of retroflexion and its generalizations avoids the confounding factor of structural simplicity.

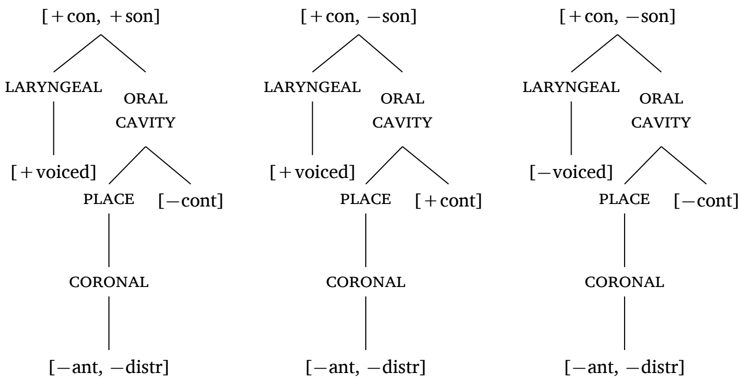

Figure 6: Structure of retroflex lateral approximants, fricatives and plosives.

The experiments dealing with retroflexion will not consider the effect of ease of articulation. The reason for this is that the statements regarding differences in the articulation depending on consonant manner are contradictory (see below). Unfortunately, it is beyond the scope of this dissertation to accurately investigate the articulatory differences in retroflex consonants of different manners. For the sake of completeness, let me briefly explain the difficulties establishing ease of production for retroflexes. Retroflex consonants are traditionally described as those consonants that are articulated with a backward curled tongue tip or tongue blade (Hamann, 2003b; Trask, 1996). According to Stausland Johnsen (2012) there is no difference in the articulation of the retroflex consonants of different manners because the process of changing the place of articulation – from alveolar to a point further back in the mouth due to the backward bending – is the same for all alveolar consonants. Hamann (2003b), however, points out that retroflex plosives and retroflex laterals have the same articulatory characteristics (Dixon, 2011; Ladefoged & Maddieson, 1996) while retroflex plosives and retroflex fricatives differ in their articulatory characteristics (Keating, 1991). Whereas retroflex plosives require a backward bending of the tongue for only a very short time during closure, retroflex fricatives require a backward bending of the tongue during the whole articulation (Hamann, 2003b; Ladefoged & Maddieson, 1996). This suggests that the retroflexion of fricatives might be of more effort than that of plosives or laterals. This assumption is based on articulatory data of retroflex consonants from Indo-Aryan, Dravidian, Indic and Australian languages. There is, however, no comment on retroflex consonants in North-Germanic languages, such as Norwegian, the language my experimental stimuli are recorded in.

To sum up, retroflexion is used as an experimental test case to investigate whether learners are biased by perceptual cues when learning and generalizing a phonological pattern. As there is the postalveolar fricative [ʃ] in German that is perceptually similar to the retroflex fricative [ʂ], this offers me the possibility to investigate in how far participants rely on language-specific perceptual cues based on the phonetic similarity to sounds in their native language and in how far they rely on language-general perceptual cues based on ease of perception. Language-general perception (ease of perception) predicts a learning advantage for fricatives and more generalizations to fricatives for all learning groups and language-specific perception (phonetic similarity to L1) predicts a learning advantage for fricatives but no generalizations to and from fricatives. For detailed predictions see chapter 4.4 and 5.

1.5 Overview of the experiments

In this chapter I will briefly summarize the experiments I conducted by focusing on their aims, their procedures and their results.

Chapter 2: Production, perception and learning of vowel nasalization In this experiment series I focused on the question of what aspects of the phonetics of a phonological pattern affect its learning. Is learning a phonological pattern more strongly biased by ease of articulation or by ease of perception? Does a phonetic learning bias influence generalizations as proposed by ease of perception and ease of articulation or is it the case that the phonetics of a phonological pattern is used by learners as proposed by phonetic similarity? To this end I considered the effects of ease of articulation, ease of perception and phonetic similarity on learning and generalizations in an artificial language learning experiment. Ease of articulation was assessed by the literature already mentioned and the data of a production experiment, ease of perception was assessed by the literature already mentioned and the data of a perception experiment and phonetic similarity was assessed by phonetic measurement of the stimuli.

For the phonetic measurement (chapter 2.1) of the stimuli formant frequencies of oral and nasal vowels of different heights were measured. The data was used to measure – based on formant frequencies – the phonetic distance between vowels of different height. This was done to establish the phonetic similarity of vowels. The results show that low and mid vowels are closer to each other than high and mid vowels. This is true for oral vowels and for nasal vowels and it is in line with the predictions based on the vowel names.

The production experiment (chapter 2.2) consisted of a reading task and an acoustic analysis of the recordings. In the production experiment German native speakers were asked to read aloud pseudowords with vowels of different height either after a non-nasal consonant (CV), after a nasal consonant (NV) or before a nasal consonant (VN). The data was used to measure – based on spectral tilt – how heavily vowels of different height are nasalized in the context of nasal consonants in comparison to vowels in the context of non-nasal consonants. This was done to establish the ease of articulation of nasal vowels. The results show that the difference in acoustic nasality between the oral and the nasal contexts is greatest in low and mid vowels, and smallest in high vowels. This is in line with former experimental results and the predictions based on anatomy.

The perception experiment (chapter 2.3) was a forced-choice identification task that tested the perception of nasal and oral vowels of different height by German native speakers. In this experiment participants were asked to identify oral and nasal vowels of different height from a choice of six alternatives. The data was used to measure – based on confusions – how perceptually similar oral and nasal vowels for each height are. This was done to establish the ease of perception of nasal vowels. The results show that oral and nasal high vowels are less often confused with each other than oral and nasal low vowels or oral and nasal mid vowels. This is in line with former experimental results and the predictions based on acoustics.

The artificial language learning experiment (chapter 2.4) consisted of a perceptual training phase and a perceptual two-alternative forced-choice test phase. In this experiment I used morphophonological vowel nasalization of different heights to investigate the role of phonetics in learning and generalizations. In the experiment German native speakers were taught an artificial language in which plurals are marked by the suffix [m] with a preceding nasalized vowel, whereas diminutives are marked by the suffix [l] without a preceding nasalized vowel. The learning results support the role of ease of perception more than the role of ease of articulation. The generalization results support the role of phonetic similarity more than the role of ease of perception and ease of articulation as the learners generalized more to a phonetically similar pattern than to a perceptually or articulatory easy pattern.

Chapter 3: Articulation and learning of vowel nasalization This experiment aimed at testing whether the articulation of nasal vowels is modified by German native speakers due to auditory exposure to vowel nasalization of different height. Do speakers increase the degree of nasalization so that they are better at imitating vowel nasalization after exposure to auditorily presented vowel nasalization? If so, do speakers imitate the nasalization of the articulatory easy pattern, of the perceptually easy pattern, of the phonetically similar pattern or of the phonetically less frequent pattern better? To this end I considered the effects of ease of articulation, ease of perception, phonetic similarity and perceptual frequency on imitations in a reading task incorporated in an artificial language learning experiment.

The artificial language learning experiment (chapter 3.2) consisted of a first reading task, a perceptual exposure phase and a second reading task. In the two reading tasks German native speakers read aloud sentences that contain French loan words with nasal vowels. In the perceptual exposure phase participants were taught the same artificial language as in the first experiment series described above. The recordings of the French loan words before and after exposure to vowel nasalization were used to measure – based on acoustic analysis – how auditory exposure to vowel nasalization of different height increases the degree of articulatory vowel nasalization. I measured spectral tilt and the degree of vowel space centralization by looking at vowel quality changes and calculating the vowel space area. The results were quite variable so that none of the predictions could be entirely confirmed. However, the results suggest a greater increase of nasalization in the trained height by participants being exposed to high vowel nasalization – that is the perceptually easy and the phonetically less frequent pattern. As participants did not generalize their learned pattern in a systematic way, I was not able to make any clear statement about how phonetics affects the generalizations in production. The results suggest, however, that speakers generalize to phonetically similar patterns.

Chapter 4: Perception and learning of retroflexion This experiment series focused on the question to what extent language-general perceptual knowledge and native language knowledge bias phonological learning and generalizations. Does a phonetic bias influence learning and generalizations as proposed by ease of perception or is it the case that the phonetics of a phonological pattern that is compared to the phonetics of a native phonological pattern is used by learners as proposed by the perceptual similarity to an L1-pattern? To this end I considered the effects of ease perception and perceptual similarity to native patterns on learning and generalizations in an artificial language learning experiment. Ease of perception and perceptual similarity were assessed separately by the data of two perception experiments.

The first perception experiment (chapter 4.2) was a forced-choice identification task that tested the perception of alveolar and retroflex consonants of different manner by German native speakers. In this experiment participants were asked to identify alveolar and retroflex consonants of different manner from a choice of ten alternatives. The data was used to measure – based on confusions – how perceptually similar alveolar and retroflex consonants for each manner are. This was done to establish the ease of perception of retroflex consonants. The results show that alveolar and retroflex fricatives are less often confused with each other than alveolar and retroflex plosives, nasals and lateral approximants. This is in line with former experimental results and the predictions based on acoustics.

The second perception experiment (chapter 4.3) was a forced-choice discrimination task that tested the discriminability of non-native retroflex consonants and native (post)alveolar consonants by German native speakers. In this experiment participants were asked to decide whether the first or the second syllable was repeated in a triplet. The data was used to measure – based on percentages correct – whether the participants can discriminate the native and non-native consonants. This was done to establish the perceptual similarity between non-native and native consonants. Special focus lay on the discrimination of the retroflex and postalveolar fricative. The results show that German native speakers can only poorly discriminate the retroflex and postalveolar fricative, whereas they performed much better on the discrimination of alveolar and retroflex plosives, lateral approximants and fricatives.

The artificial language learning experiment (chapter 4.4) consisted of a perceptual training phase and a perceptual two-alternative forced-choice test phase. In this experiment I used morphophonological retroflexion of consonants of different manner to investigate the role of phonetics in learning and generalizations. In the experiment German native speakers were taught an artificial language in which plurals are marked by the suffix [ɾ] with a preceding retroflex consonant, whereas diminutives are marked by the suffix [n] without a preceding retroflex consonant. The results support the role of language-specific perception more than the role of language-general ease of perception as the learners misperceived the retroflex fricative as postalveolar fricative so that no generalizations from or to the retroflex fricative appeared.

Chapter 5: Learning morphophonology without alternation
In this chapter I investigated whether a morphophonological pattern can be generalized without evidence for a morphophonological alternation and whether this lack of reliability about an alternation forces participants to rely more on language-general (phonetically grounded) perceptual information than on language-specific (L1-based) perceptual information when learning and generalizing a pattern, for which both kinds of information are available, namely retroflexion. Previous studies

suggested that participants in less reliable situations, e.g. short training phases and training input with exceptions, are more strongly affected by phonetics and less strongly affected by L1 than participants in more reliable situations, e.g. long training phases and training input with fewer exceptions. To this end, I conducted an artificial language learning experiment without evidence for an alternation during training. This training phase contrasts with the training phase in the previous artificial language learning experiment on retroflexion in which evidence for an alternation was provided.

The artificial language learning experiment (chapter 5.2) consisted of a perceptual training phase and a perceptual two-alternative forced-choice test phase as the artificial language learning experiment described above. In the experiment German native speakers were taught an artificial language in which plurals are marked by a retroflex consonants following [ɾ]. However, no singular forms were presented during training, which on the one hand made training shorter and on the other hand failed to provide direct evidence for a morphophonological alternation. If singulars, plurals and diminutives are presented, and if retroflexion occurs only in plurals, learners have evidence for an alternation in the plurals based on the singular. However, if only plurals and diminutives are presented, and if retroflexion occurs only in plurals, learners have no evidence for a phonological alternation, since it can be just another consonant in those items. The learning results were similar to the ones in the previous artificial language learning experiment but the performances with the untrained items were much weaker and offered more support for the role of language-general ease of perception than for the role of language-specific perception. This is in line with the prediction that the phonetic bias is stronger, the less reliable the information about the morphophonological pattern is. As the learners did not generalize their learned phonological pattern, evidence for a morphophonological alternation – knowledge about the relation of forms in a paradigm – is necessary to accurately predict novel items based on the learned pattern.

2 Production, perception and learning of vowel nasalization

2.1 Introduction

The aim of the present study is to directly address the questions whether learning a morphophonological pattern is biased by ease of articulation or ease of perception (Finley, 2017; Finley & Badecker, 2007, 2009; Wilson, 2003, 2006), and whether learners are biased by the ease of articulation, the ease of perception or the phonetic similarity (Kapatsinski, 2013a) when generalizing this learned morphophonological pattern. I further aim at investigating to what extent markedness can be used to predict learning biases based on ease of perception or ease of articulation.

To this aim I conducted a learning experiment in which adult German native speakers were taught an artificial language consisting of CVCV singulars, CVCṼ[m] plurals and CVCV[l] diminutives. The final vowel of the singular was either high [i], mid [ɛ] or low [a]. These different vowel heights allowed me to test whether learners are biased by phonetics and if so, whether they rely more on ease of perception, ease of articulation or phonetic similarity as will be elaborated further below.

I investigated the similarity of the vowels by measuring their formants and calculating their distance, I assessed the ease of perception by means of a perception experiment (see chapter 2.3) and I assessed the ease of articulation by means of a production experiment and an acoustic analysis (see chapter 2.2). I further assessed ease of perception and ease of articulation by investigating the typology of nasal vowels.

According to the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006) the predispositions based on ease of perception and ease of articulation are encoded as markedness in our grammars. In the case of vowel nasalization, low nasal vowels are easiest to articulate because the muscle that lowers the tongue body simultaneously lowers the velum, thereby allowing air to flow through the nose (Bell-Berti, 1993; Ohala, 1975; Zsiga, 2013), whereas high nasal vowels are easiest to perceive since even slight lowering of the velum has a large spectral effect (Ohala, 1975). This is in line with Hajek (1997) and Hajek & Maeda (2000) who claims that there are two universal factors that affect the emergence of contrastive vowel nasalization: a preference for low vowel nasalization and a preference for high vowel nasalization. Hajek & Maeda (2000), however, claim that both preferences are due to perception. They argue that low nasal vowels are favored perceptually because low vowels have an intrinsically longer duration which makes the nasalization more salient (Laver, 1994; Lehiste, 1970).

To assess these reported effects of ease of articulation and ease of perception independently, I investigated the distribution of the height of nasal vowels in the languages of the world. Since ease of articulation favors low nasal vowels and ease of perception favors high nasal vowels, I will expect that there will be relatively few languages with nasal mid vowels, if these factors affect the typology of nasal vowels due to markedness.

I did this by creating a dataset of 98 languages that are all languages with short nasal vowels. I selected languages with short nasal vowels because I used these short vowels in my artificial language learning experiment. The dataset was drawn from the 451 languages of the UPSID corpus (Maddieson, 1984). I used the interface programmed by Henning Reetz to do this: http://web.phonetik.uni-frankfurt.de/upsid.html. The UPSID corpus contains a representative set of languages from all language families. I then annotated for each language with short nasal vowels which vowel heights were present. The typology of nasal vowels matches their ease of articulation and ease of perception (Hajek & Maeda, 2000; Ohala, 1975; Zsiga, 2013). Nasal mid vowels are indeed infrequent in comparison to nasal low and high vowels. In only 47% (n = 46) of the languages there are mid nasal vowels, whereas 90% (n = 88) have high nasal vowels and 93% (n = 91) have low nasal vowels (see Table 1).

Table 1: Distribution of vowel height in languages with short monophthong vowels.

Vowel height	**Ṽ**	**V**	**Ṽ among V**
high	88	416	21%
mid	46	240	19%
low	91	432	21%

However, one has to keep in mind that mid vowels are generally rare in the languages of the world. Thus, these numbers may be misleading. To check the percentages of nasal vowels among the total number of vowels of different heights I created a second dataset. This second dataset contained all 449 languages that have short monophthong vowels of the heights high, mid or low – independent of any spectral modification. The results can be seen in Table 1 as well. It turns out, that mid vowels are indeed infrequent in comparison to low and high vowels. In only 53% (n = 240) of the languages there are mid vowels, whereas 93% (n = 416) have high vowels and 96% (n = 432) have low vowels.

When comparing the distribution of short nasal monophthong vowels with the distribution of all short monophthong vowels in Table 1 one can observe that there is the same amount of high and low oral and nasal vowels (about 90%) and a smaller amount of mid oral and nasal vowels (about 50%). Thus, the percentage of nasal vowels among all vowels should not differ across height.

By calculating the percentages of nasal vowels among all vowels, it turns out that the distribution of nasal vowels does indeed not differ depending on height: 21% of all high vowels are nasal, 19% of all mid vowels are nasal and 21% of all low vowels are nasal. It thus seems as if the low number of languages with mid nasal vowels is due to the low number of languages with mid vowels in general.

As a consequence, the proposed encoding of the predispositions as markedness by the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006) does not hold for vowel nasalization when considering the percentages of nasal vowels among all vowels but only when considering the raw numbers. This is an important finding as it questions the encoding of a phonetic bias in markedness, one of the core statements of the framework of phonetically based phonology.

Independent of the potential failure of a phonetic bias to be encoded in markedness, a phonetic bias might be present during learning. To test this, I did a production experiment with an acoustic analysis to check whether low nasal vowels are more heavily nasalized than high nasal vowels by German native speakers (see chapter 2.2) and a perception experiment to check whether high nasal vowels are easier to perceive than low nasal vowels by German native speakers (see chapter 2.3).

To further test for the phonetic similarity as a factor influencing the learning behavior of the participants, it is necessary to check whether (lower) mid [ɛ] is phonetically more similar to low [a] than to high [i] in the stimuli I used in the artificial language learning experiment. Foreshadowing a more detailed description of my items in chapter 2.4.1, I provide data on the phonetic properties of the last vowel in the learning experiment items, which I call *target vowel* here. There are three target vowels: [i], [ɛ] and [a]. In the training phase these were the last vowels in the items and in the plural they were followed by an [m]. I measured the first and second formants at the mid point of these vowels by means of a Praat script (Boersma & Weenink, 2019). Figure 7 provides a graphic overview of the F1- and F2-values of the target vowels. The boxed labels in the graphic are the mean values calculated with R (R Core Team, 2015) and plotted with ggplot2 (Wickham, 2009).

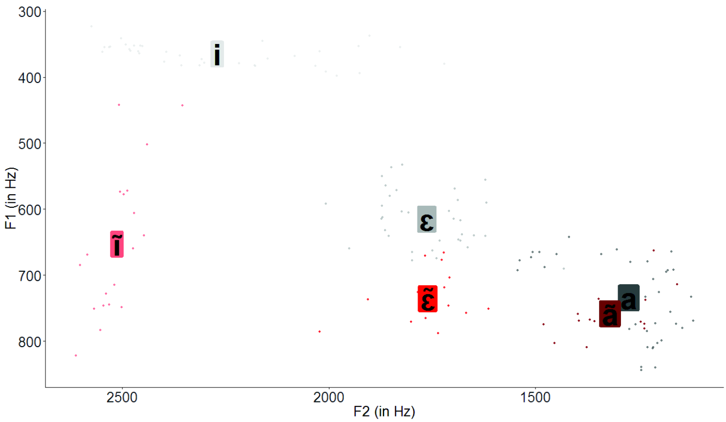

Figure 7: Acoustic properties of the target vowels in the items of the artificial language learning experiment on vowel nasalization.

The graph shows that [ɛ] and [a] are closer together than [i] and [ɛ], or [i] and [a] and that [i] is closer to [ɛ] than to [a]. This is true for the oral vowels and for the nasal vowels. This means that [ɛ] and [a] resemble each other more than [i] resembles [ɛ], or [i] resembles [a] and that [i] resembles [ɛ] more than [a].

The visual impression of Figure 7 is confirmed by the acoustic distance I calculated for these vowels. I estimated their distance by considering their mean values as points on a two dimensional plane, and I then used Pythagoras' theorem to calculate the distance between these points. I used the formula in (1).

(1) $$\sqrt{(\mathrm{F2}_\alpha - \mathrm{F2}_\beta)^2 + (\mathrm{F1}_\alpha - \mathrm{F1}_\beta)^2}$$

The results of these calculations are presented in Table 2. The smaller the number, the closer the vowels are. The smallest distance is between [a] and [ɛ], followed by [i] and [ɛ], while the greatest distance is between [i] and [a]. This is true for the oral vowels as well as for the nasal vowels. Thus, the phonetic similarity based on the vowel names and their position in the general vowel chart is consistent with the phonetic similarity in the experimental items.

Table 2: Acoustic distance (in Hz) between target vowels in the items of the artificial language learning experiment on vowel nasalization.

Vowel pairs	**Distance**	
	oral	**nasal**
[i] & [ɛ]	565	757
[i] & [a]	1063	1199
[ɛ] & [a]	504	443

I will now continue with experiments testing the articulation and the perception as well as the learning and generalization behavior of German native speakers.

2.2 Production experiment and acoustic analysis

According to Laeufer (2010) and Wiese (1996) there are no phonemic nasal vowels in German and even allophonic nasalization is most often restricted to French loan words and to the speech of some speakers from southern areas of Germany. As far as I know, there is no systematic research on the production of coarticulatory nasalized vowels in German. The research is most often limited to the pronunciation of nasal vowels in French loan words by German native speakers (Greisbach, 2003; Laeufer, 2010). To investigate whether there is coarticulatory vowel nasalization in Northern German – a variety in which neither phonemic nor allophonic nasal vowels are present according to Laeufer (2010) and Wiese (1996) – I asked participants from the region in the western part of North Rhine-Westphalia to read aloud pseudowords with vowels in the context of nasal and non-nasal consonants. People who live there usually pronounce nasal vowels as an oral vowel followed by [ŋ]: The French loan word *Croissant* is pronounced as [kro.sɔŋ]. Sometimes *Croissant* is even incorrectly written as *Crossong* illustrating the 'non-nasalized German' pronunciation of originally nasal vowels (seen at a bakery at a train station in the western part of North Rhine-Westphalia). Nevertheless, it might be the case that in Northern German allophonic nasalization is present due to automatic assimilatory processes driven by coarticulation (Bell-Berti, 1993; Crystal, 2008). As the movement of the velum from its closed position to its open position that is large enough to produce perceptible nasalization, and vice versa, takes about 50 ms, a certain amount of coarticulation is physiologically necessary (Ohala, 1975).

To test this I did a production study and I analyzed the degree of nasalization of vowels of different heights and in different contexts. I wanted to establish whether adult native speakers of Northern German nasalize vowels in the context of nasal consonants and if so, whether they produce stronger nasalization in low nasal vowels than in high nasal vowels as would be expected based on ease of articulation. As Northern Germans are said to produce no vowel nasalization, they are appropriate participants to investigate the ease of articulation. If I tested Southern Germans who are said to sometimes nasalize vowels, their individual experience with nasal vowels would have been a confounding factor. Moreover, it might have been the case – as nasalization is familiar to them – that the degree of nasalization would not differ depending on vowel height.

During the articulation of nasal vowels the velum is lowered and as a result the nasal cavity then acts as an additional resonator (Ohala, 1975). The most prominent acoustic consequences of having two resonators – oral cavity and nasal cavity – are the reduction of the first formant's amplitude in addition to other acoustic effects (House & Stevens, 1956; Schwartz, 1968) (for details chapter 1.3). Measuring nasality is thus possible by looking at changes in amplitude in the spectrum. I measured the A1−P0 difference as this value was used in several other studies (Chen, 1995, 1997; Styler, 2015; Zellou, Scarborough & Nielsen, 2016; Zellou et al., 2017) and revealed congruent and robust results (Styler, 2017). A1 refers to the amplitude of the highest harmonic near the first formant whereas P0 refers to the amplitude of a nasal peak at about 250 Hz. As proposed by Chen (1997) A1 is either the first (H1) or the second harmonic (H2), which depends on whichever of the two harmonics has a higher amplitude. As A1 and P0 can occur at the same place in high vowels, some researchers measure the A1−P1 difference in high vowels (Chen, 1997; Styler, 2015). P1 refers to the amplitude of a nasal peak harmonic at about 950 Hz closest to the first formant. However, Styler (2017) considered the A1−P0 difference as the most robust measurement for vowels of all heights. The smaller the A1−P0 difference, the higher the amount of nasality. The reason behind this is that a larger velopharyngeal port opening causes a more prominent nasal peak (P0 (and P1)) as well as a more reduced amplitude of the first formant (A1) (Chen, 1997).

For the present experiment I predicted that if German native speakers benefit from a language-independent phonetic bias based on articulatory ease, they will nasalize low vowels to a greater degree than non-low vowels in the context of nasal consonants. If my participants do not produce a different degree of nasalization on low and non-low vowels in the context of nasal consonants, they will not benefit from articulatory ease.

2.2.1 Method

Stimuli I created a list of 75 pseudowords with legal German phonotactics (Wiese, 1996). The stimuli consisted of 25

$C_1V_1C_2V_2$-, 25 $C_1V_1C_2V_2$[m]-, and 25 C_1V_1[m]V_2-items. The consonants [p], [d], [k], [ʃ] and [v] were used in C_1-position and the consonants [b], [t], [g], [s] and [f] were used in C_2-position. V_1 was one of the vowels [o] and [u] and V_2 was one of the vowels [a], [ɛ], [i], [o] and [u]. V_2 was the vowel in which the nasality was measured. I measured the nasality in a vowel after a non-nasal consonant (context CV), before a nasal consonant (context VN), and after a nasal consonant (context NV). As the participants were not familiar with IPA transcriptions I gave all stimuli in German orthography. An example of a CVCV-item is <wusa> [vusa], an example of a CVCV[m]-item is <wusam> [vusam], and an example of a CV[m]V-item is <wuma> [vuma]. A complete list of all stimuli used in this experiment is given in the appendix (see Table 26).

Procedure The participants read aloud the stimuli that were embedded in one of two target sentences. Vowel-final items (CVCV and CV[m]V) were embedded in *Ich habe X gesagt.* 'I said X.' and nasal-final items (CVCV[m]) were embedded in *Ich habe X erblickt.* 'I saw X.'. The sentences were displayed separately on a computer screen of a Windows laptop in randomized order. After having read out one sentence the participants pressed a button and the next sentence was displayed.

The recordings took place in an anechoic booth in the phonetics laboratory of the Heinrich-Heine University Düsseldorf to avoid disturbing noise. For the recordings a microphone (phantom power 48V), a Sound Devices amplifier, a Marantz recorder and a Transtec computer were used. The sampling rate was 48 kHz.

Participants 15 adult native speakers of Northern German (9 identified as female, 6 identified as male, mean age: 23.7, range: 21-28) were recorded. No one reported knowledge of a language that uses nasalized vowels distinctively. All participants had normal or corrected-to-normal vision, no reported hearing problems and did not suffer from hoarseness during the recordings. They participated voluntarily. Prior to the experiment each participant filled out a consent form and a questionnaire with which data about the participants' age, sex, native language, foreign language skills, the region they had been grown up and education was collected.

2.2.2 Results

Prior to analysis the data had to be prepared, which means that textgrids were created in Praat (Boersma & Weenink, 2019), in which the items as well as the vowel under investigation were labeled. The selection and the labeling of the vowel had to be consistent across all recordings so that the measurements were not confounded by inconsistent vowel boundaries. The vowels' starting point and the vowels' end point was always a zero-crossing point. A zero-crossing point is a time point in the oscillogram at which the amplitude is zero, which means complete silence. The vowel boundaries were set based on visual inspections of the wave form and the spectrogram. The first quasiperiodic wave of the vowel was chosen as starting point and the last quasiperiodic wave was chosen as end point of the vowel. This selection was in agreement with the visible formants in the spectrogram. Figure 8 shows as an example the annotation made by means of a textgrid in Praat (Boersma & Weenink, 2019). Oscillogram (above) and spectrogram (below) of [pogam] (speaker-01: male) are shown. The stimulus item [pogam] and the vowel [a] were selected and labeled in the textgrid tiers 1 and 2 below.

The original data set consisted of 1125 vowels (75 words $\times$ 15 participants). Eleven words had to be excluded due to mispronunciations. As nasality has similar acoustic features as creaky voice (Zhang, 2015), I also had to exclude all words (n = 36) produced with creaky voice. Creak has several acoustic properties, e.g. irregular F0 (measurable by harmonicity-to-noise-ratio (HNR)), low F0, glottal constriction (measurable by spectral tilt) and damped pulses (Keating, Garellek & Kreiman, 2015). The literature provides no guidance as to the cut-off point between creaky voice and nasality. I therefore conducted the following procedure: I decided to measure HNR with Praat (Boersma & Weenink, 2019) to determine which of the recordings sound creaky. As sounds with lower HNR are creakier, I ordered the stimuli according to their HNR-value. A phonetically trained linguist and I listened independently from each other to the ones with the lowest HNR. In addition to that we checked the distance between the pulses visually because irregularly spaced pulses are a further hint to creaky voice. By doing so, both listeners labeled all sounds with an HNR lower than 8 dB as creaky, which was then chosen as cut-off point. In Figures 15 and 16 in chapter 3.2.2 you can see the difference between recordings with creaky voice (Figure 15) and without creaky voice (Figure 16).

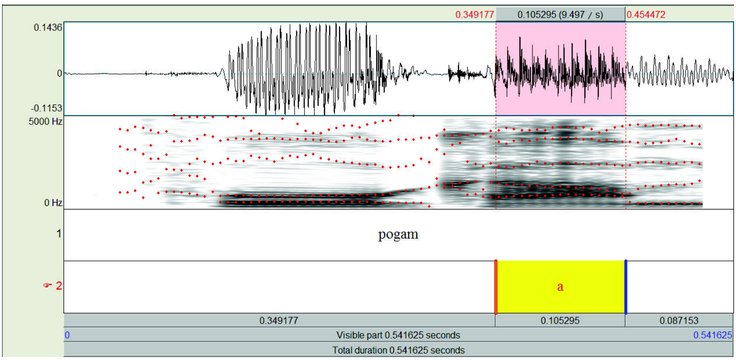

Figure 8: Example of the annotation made by means of a textgrid.

The remaining vowels (n = 1078) were further analyzed. I measured nasality acoustically by means of the A1−P0 value to objectively compare the amount of acoustic nasality in the items depending on vowel height. The lower the A1−P0 value, the greater the amount of nasality. With Praat (Boersma & Weenink, 2019) and the Nasality Automeasure script (Styler &

Scarborough, 2017) I measured A1−P0 values at five different time points for each vowel in a word. These time points related to the beginning, to 25%, to 50%, to 75% and to the end of the vowel portion. Due to problems with the measurement indicated by the script (A1 = P0 (n = 1362), Shallow (n = 268), LoPitch (n = 124), HighF1 (n = 44), HiPitch (n = 29), HarmDev (n = 27)), only 3536 data points of the original 5390 data points (1078 vowels × 5 time points) were analyzable. These errors occur when no A1−P0 value can be measured because both – A1 and P0 – are the same (A1 = P0), when H1 and H2 cannot be detected on the side of a single peak (Shallow), when H1 is below 80 Hz (LoPitch), when F1 is above 1000 Hz (HighF1), when H2 is above 300 Hz (HiPitch) or when H2 is more than half an F0 above or below 2 × H1 (HarmDev) (see Styler & Scarborough (2017) for details). These errors – especially the error A1 = P0 – and thus the failure to measure nasality might be a first hint at the results: There might only be a small amount of nasality in the items, if at all.

For the analysis I calculated the mean A1−P0 value for each context depending on only three of the five measured time points. For VN-items these time points related to 50%, 75% and to the end of the vowel interval and for NV-items these time points related to the beginning, to 25% and to 50% of the vowel interval. This was done because I expected the greatest amount of nasality in the vowel portion next to the nasal consonant. As the A1−P0 values of NV- and VN-items were compared to those of CV-items, A1−P0 values for CV-items were measured at time points relating to 50%, 75% and to the end of the vowel interval for the comparison with VN-items and at times points relating to the beginning, to 25% and to 50% of the vowel interval for the comparison with NV-items.

I was interested in whether there are differences in the degree of nasalization in vowels of different height. I therefore measured the mean difference between the A1−P0 values in the oral context (CV) and the nasal contexts (VN and NV) for each vowel height separately. The greater this difference between oral and nasal contexts is, the more nasal the vowels in the nasal contexts are. Additionally, I calculated for each vowel height (high, mid and low) in each comparison (CV-VN and CV-NV) a separate linear mixed-effects model in R (R Core Team, 2015) with the R packages lme4 (Bates, Mächler, Bolker & Walker, 2015) and lmerTest (Kuznetsova, Brockhoff & Christensen, 2018). In the models which fitted the data best (as assessed by backward stepwise elimination (Baayen, 2008)) the A1−P0 VALUE was the independent variable and the CONTEXT (oral (CV) and nasal (NV or VN)) was the dependent variable. PARTICIPANT and ITEM served as random intercepts (R syntax: lmer($\mathrm{A1{-}P0} \sim \textsc{context} + (1|\textsc{participant}) + (1|\textsc{item})$)). With the R package (R Core Team, 2015) ggplot2 (Wickham, 2009) and Rmisc (Hope, 2013) I plotted the mean A1−P0 values depending on vowel height (high, mid and low) and context (CV, VN and NV). In Figures 9 and 10 the A1−P0 values (in dB) $\pm$ 1.96 SE are shown for high, mid and low vowels in the oral context (gray) and in the nasal context (red). Higher bars (= lower values) correspond to greater acoustic nasality. Note that the y-axis is turned upside down so that negative values that correspond to a greater amount of nasality are illustrated with raising bars.

The three linear mixed-effects models – one for each vowel height – for the comparison of CV and NV show, that there was no significant difference (−0.62 dB) between CV (−0.74 dB) and NV (−1.36 dB) for high vowels (Est. = −0.8226, SE = 0.5717, df = 11.449, t = −1.439, p = 0.177). There was, however, a greater amount of nasality (a lower A1−P0 value) in NV (−2.24 dB) than in CV (2.02 dB) in mid vowels. This difference (−4.26 dB) was significant (Est. = −4.2332, SE = 0.9439, df = 18.539, t = −4.484, $p < 0.001$). The same is true for low vowels. There was a significant greater amount of nasality (a lower A1−P0 value of −3.94 dB) in NV (−1.92 dB) than in CV (3.02 dB) in low vowels (Est. = −4.8258, SE = 0.5628, df =328.6847, t = −8.575, $p < 0.001$). An overview of the results for the CV-NV-comparison can be found in Figure 9.

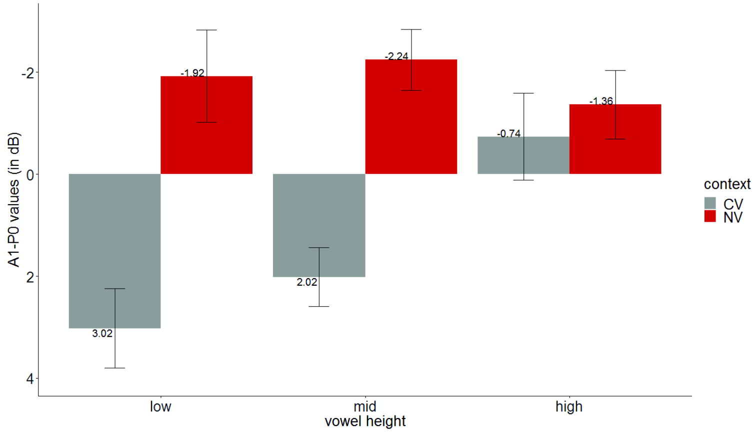

Figure 9: A1−P0 values (in dB) ± 1.96 SE for high, mid and low vowels in the oral context CV (gray) and in the nasal context NV (red).

The three linear mixed-effects models – one for each vowel height – for the comparison of CV and VN show similar results. There was only a marginally significant difference (0.88 dB) between CV (0.22 dB) and VN (1.10 dB) for high vowels (Est. = 1.4081, SE = 0.7094, df = 17.220, t = 1.985, p = 0.063). There was, however a greater amount of nasality (a lower A1−P0 value) in VN (0.33 dB) than in CV (2.18 dB) in mid vowels. This difference (−1.85 dB) was significant (Est. = −1.8369, SE = 0.5018, df = 17.919, t = −3.661, $p < 0.01$). The same is true for low vowels. There was a significant greater amount of nasality (a lower A1−P0 value of −1.40 dB) in VN (0.86 dB) than in CV (2.26 dB) in low vowels (Est. = −1.3845, SE = 0.5345, df = 332.791, t = −2.590, $p < 0.05$). An overview of the results for the CV-VN-comparison can be found in Figure 10.

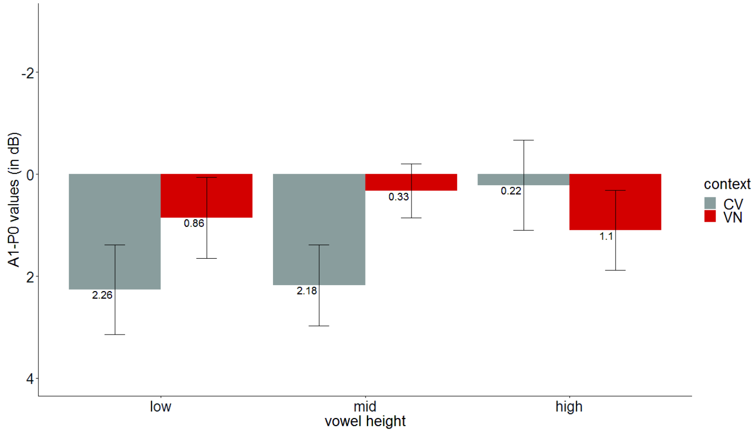

Figure 10: A1 − P0 values (in dB) ± 1.96 SE for high, mid and low vowels in the oral context CV (gray) and in the nasal context VN (red).

2.2.3 Discussion

As there has not been any systematic investigation of coarticulatory vowel nasalization in Northern German yet, the gained results offer new findings with which one can evaluate the restrictions of coarticulatory nasalization – concerning vowel height, position and degree of nasalization – in a language that is generally considered to have neither phonemic nor allophonic nasalization. The results show that there is coarticulatory nasalization in Northern German. Based on the A1 − P0 values non-high vowels are coarticulatory nasalized in NV and in VN contexts by Northern German native speakers. High vowels are not nasalized in NV or VN contexts. This means that there is coarticulatory vowel nasalization in non-high vowels in Northern German. Thus German participants benefit from articulatory ease.

To rank the degree of nasality, one has to compare the measured A1 − P0 with those from other languages. Chen (1997) reported mean differences of A1 − P0 values in oral and nasal contexts of 6-8 dB for English and 3-9 dB for French, the mean differences of A1−P0 values in the Northern German data in oral and nasal contexts are the following: mid vowels$_{\text{CV-VN}}$: 2 dB, mid vowels$_{\text{CV-NV}}$: 4 dB, low vowels$_{\text{CV-VN}}$: 1 dB, low vowels$_{\text{CV-NV}}$: 4 dB. This means that there is only a weak amount of nasality in non-high vowels in Northern German in comparison to other languages with allophonic nasalization such as English or with phonemic nasalization such as French. The differences between oral and nasal contexts in non-high vowels are greater in NV- than in VN-context – pointing at a stronger progressive nasalization and a weaker regressive nasalization.

Northern German native speakers are guided by phonetics when articulating unfamiliar sound patterns as for example nasal vowels of different height. This is in line with the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006) and gives evidence for the ease of articulation of low nasal vowels. As the literature gives no information as to how easy the nasalization of mid vowels is, I cannot state whether my data of mid vowels is in line with the proposed ease of articulation. According to the data of Northern Germans it seems as if the articulation of mid nasal vowels is as easy as the nasalization of low nasal vowels. Whether this is due to the same anatomic factors or due to other factors, cannot be assessed with the present analysis.

Nevertheless, as shown in Figure 7 and Table 2, the mid vowel [ɛ] is a lower mid vowel that is acoustically more similar to the low vowel [a] than to the high vowel [i]. As articulatory differences have most often direct acoustic consequences – for vowels these consequences are manifested in frequency changes of F1 and F2 – the nasalization of lower mid [ɛ] seems – due to the acoustic proximity to the easy to nasalize [a] – to be of less effort than that of a high vowel. Note that I did not analyze the acoustic distance between [a] and [o]. However, as [o] is traditionally described as upper mid vowel, I suggest [o] to be acoustically and articulatory more distant from [a] than from [ɛ]. The combined analysis in which [ɛ] and [o] are summarized under mid vowels thus may suffer from the different acoustic distances between [a] and both mid vowels. A closer look in the data confirms this: [ɛ] is more heavily nasalized in NV (−4.22 dB) and VN (−0.36 dB) than [o] (NV: 0.14 dB, VN: 1.06 dB), and especially for NV the difference to the oral context (CV) is much greater for [ɛ] (−5.78 dB) than for [o] (−2.68 dB), whereas the difference between CV and VN is quite similar ([ɛ]: −1.93 dB, [o]: −1.68 dB). Based on this short analysis, I suggest that the great amount of nasality in mid vowels is mainly due to the data of the vowel [ɛ] and that lower mid vowels benefit more from ease of articulation (like low vowels), whereas upper mid vowels benefit less from ease of articulation. Upper mid vowels, however, may benefit more from ease of perception (like high nasal vowels). To investigate this gradience is beyond the scope of this dissertation so that in the following experiments only [a], [ɛ] and [i] are used in the stimuli. Independent of these differences, as Northern German native speakers, who should not be familiar with vowel nasalization at all, are influenced by ease of articulation in their productions, I expect that this is true for all German native speakers.

2.3 Perception experiment

Similar to the production of vowel nasalization by German native speakers, there is no systematic research on the perception of vowel nasalization by German native speakers. Thus, I conducted a perception experiment to investigate how German native speakers who are not familiar with strong nasalization perceive nasal vowels of different height. In this perception experiment the participants were asked to identify oral and nasal vowels in order to assess their perceptual similarity. This task tested the perceptual confusability of oral and nasal vowels of different heights.

On the basis of acoustics and previous studies on the perception of nasal vowels by native speakers of American English (Bond, 1976; House & Stevens, 1956), I predicted the following for the present experiment: If German native speakers benefit from a language-independent phonetic bias based on perceptual ease, they will confuse oral and nasal high vowels less often with each other than oral and nasal non-high vowels. If my participants do not confuse oral and nasal vowels differently depending on vowel height, they will not benefit from perceptual ease.

2.3.1 Method

Stimuli The stimuli were the three oral vowels [a], [ɛ] and [i] and their three nasal counterparts [ã], [ɛ̃] and [ĩ]. In addition to the experimental vowels I used the four vowels [o], [õ], [u] and [ũ] for a short practice phase at the beginning of the experiment. There was exactly one token of each vowel for all parts of the experiment.

The vowels were spliced out of CV(C)-syllables: The oral vowels were spliced out of CV-syllables and the nasal vowels were spliced out of CV[m]-syllables. Those syllables were the last syllables of the stimuli recorded for the subsequent artificial language learning experiment (see chapter 2.4) – either of a singular form (the oral vowels) or of a plural form (the nasal vowels). The stimuli for the practice vowels were recorded separately. The consonant preceding the vowel was the same in all syllables. All vowels were sliced out of their context at the nearest zero-crossing point using Praat (Boersma & Weenink, 2019) to avoid irritating noise at the beginning or at the end of the vowels. As vowels show strong formants and a quasiperiodic wave form I chose the first quasiperiodic wave of the vowel as starting point of the vowel and the last quasiperiodic wave as end point of the vowel. This selection was in agreement with the visible formants in the spectrogram.

The stimuli were recorded in an anechoic booth in the phonetics laboratory at Heinrich-Heine University Düsseldorf to avoid disturbing noise. For the recordings a microphone (phantom power 48V), a Sound Devices amplifier, a Marantz recorder and a Transtec computer were used. The sampling rate of the recording was 48 kHz. The material was recorded by a female fully bilingual German-Portuguese native speaker who was able to produce the nasal vowels as well as the oral vowels easily. The intensity was scaled to 70 dB using Praat (Boersma & Weenink, 2019). Details about the recording of these stimuli, e.g. about the target sentence, can be found in the stimuli description of the following artificial language learning experiment on vowel nasalization in chapter 2.4.1 as the stimuli used in the perception experiment were sliced out of the stimuli recorded for the artificial language learning experiment.

Procedure The experiment was a forced-choice identification task and it was run as an online experiment using the software Percy (Draxler, 2014). Participants were asked to wear headphones and to run the experiment in a quiet place to avoid disturbing noise. The experiment lasted about 5 minutes.

In a short introductory phase they became familiar with the experiment and its setting. During this introductory phase they learned how the sounds are represented orthographically on the screen. I used the four vowels [o], [õ], [u] and [ũ] whereas oral vowels were presented in red and nasal vowels in blue with a tilde on top, for example <õ>. The participants listened to each of these vowels once. At the same time they saw a transcription on the screen. In the subsequent test phase all vowels were transcribed orthographically in the same way: Oral vowels were transcribed in red and in their standard orthography and nasal vowels in blue with a tilde on top, e.g. [a] was transcribed as <a>, [ɛ] as <ä>, and [i] as <i>. Participants listened to vowels and were forced to identify each vowel as one of these six vowels: [a], [ã], [ɛ], [ɛ̃], [i], [ĩ]. The participants responded by clicking on a box surrounding one of the six vowels on the screen. After they had made their choice the next vowel was presented. Each vowel was presented ten times in random order. The experiment lasted about ten minutes. At the end of the experiment the participants were asked to rate the difficulty of the task on a scale from 1 (very easy) to 5 (very difficult).

Participants I tested 38 adult native speakers of German (28 identified as female, 9 identified as male and 1 identified as gender diverse, mean age: 31.2, range: 18-63). No one reported knowledge of a language that uses nasalized vowels distinctively. All participants had normal or corrected-to-normal vision and no hearing problems. None of them had participated in the previous production experiment. All of them participated voluntarily. Each participant agreed to let me use the collected data and filled out a form giving information about the participants' age, sex, native language, foreign language skills and the region they had been grown up.

2.3.2 Results

As I was not able to control that every participant completed the online perception experiment, only 2032 data points of 2280 (38 participants $\times$ 60 stimuli) were available and analyzable. 248 data points were missing because 6 participants voluntarily had stopped the experiment before it was completed. The participants mean rating of the difficulty of the experiment was 2.6/5, which means that the task was neither too difficult nor to easy for the participants who have no experience with nasal vowels.

The results of the perception experiment are provided in Table 3. It is shown how often each stimulus (rows) was identified as one of the six vowels (columns). Percentages are given in brackets. Shaded cells are the ones were participants had the most misperceptions. For example, the auditory stimulus [ã] was identified as <ã> 189 times – 57% of all responses given to the auditory stimulus [ã] – and it was identified as <a> 134 times – 40% of all responses given to the stimulus [ã]. First, one can see, that oral vowels were easy to identify for German native speakers with percentages of correct responses between 83% and 94%. Second, nasal [ĩ] was more often correctly identified (84%) than nasal [ɛ̃] (64%) and nasal [ã] (57%).

Table 3: Confusion matrix of nasal & oral vowels in German.

response / stimulus	[ã]	[a]	[ɛ̃]	[ɛ]	[ĩ]	[i]
[ã]	189 (57%)	134 (40%)	4 (1%)	4 (1%)	0 (0%)	1 (<1%)
[a]	17 (5%)	322 (94%)	1 (<1%)	2 (<1%)	0 (0%)	0 (0%)
[ɛ̃]	4 (1%)	1 (<1%)	217 (64%)	117 (34%)	0 (0%)	1 (<1%)
[ɛ]	3 (1%)	2 (1%)	53 (16%)	281 (83%)	0 (0%)	1 (<1%)
[ĩ]	1 (<1%)	0 (0%)	0 (0%)	1 (<1%)	285 (84%)	54 (16%)
[i]	0 (0%)	1 (<1%)	0 (0%)	3 (1%)	39 (12%)	294 (87%)

A statistical analysis with R (R Core Team, 2015) using the packages languageR (Baayen, 2013), lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2018) confirmed the results from the visual analysis. I ran a linear mixed-effects model on a subset of the data which only included the correct responses and the incorrect ones in which only nasalization was misperceived. This corresponds to the shaded cells in Table 3. In the model which fitted the data best (as assessed by backward stepwise elimination (Baayen, 2008)) ACCURACY (correct and incorrect) served as dependent variable and VOWEL HEIGHT (high, mid and low) as independent variable. PARTICIPANT was included as random intercept (R syntax: glmer(ACCURACY ∼ VOWEL HEIGHT + (1|PARTICIPANT), family = "binomial")). The model shows that oral and nasal high vowels were less often confused with each other than oral and nasal mid (Est. = -1.0561, SE = 0.1675, z = -6.307, $p < 0.001$) and low vowels (Est. = -0.8499, SE = 0.1693, z = -5.019, $p < 0.001$). There was no significant difference in the confusions of oral and nasal low vowels and oral and nasal mid vowels (Est. = -0.2062, SE

= 0.1488, z = −1.386, p = 0.166).

To visualize the confusions between the high, mid and low oral-nasal vowel pairs I used the perceived distance d' (dprime). I calculated d' using R (R Core Team, 2015) (R syntax: dprime.mAFC(proportion of correct responses, number of alternative choices)) and the R package psyphy (Knoblauch, 2014). The results can be seen in Figure 11. This Figure was created with the package ggplot2 (Wickham, 2009) in R (R Core Team, 2015). The perceived distance between [a] and [ã] (3.42) and between [ɛ] and [ɛ̃] (3.07) was lower than the perceived distance between [i] and [ĩ] (4.99).[3]

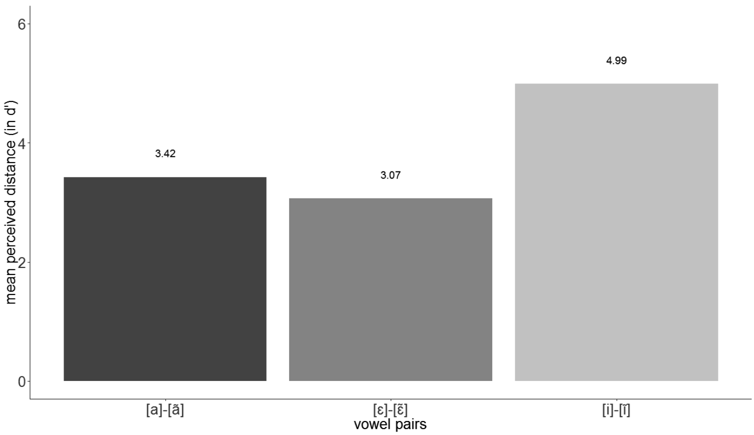

Figure 11: Perceived distance based on d' between oral and nasal vowels.

The experimental results further confirm the acoustic proximity of low and (lower) mid vowels as low and (lower) mid

[3] d'-values can range from −10 to +10, whereby positive values mean that participants can discriminate the sounds above chance.

vowels were more often confused with each other than low vowels or (upper) mid vowels with high vowels. High vowels, on the other hand, were seldom confused with other vowels, but when they were, they tended to be confused more often with (lower) mid vowels than with low vowels. When low vowels were the stimulus, (lower) mid vowels were more often given as response (11 times) than high vowels (only once). When (lower) mid vowels were the stimulus, low vowels were more often given as response (10 times) than high vowels (2 times). When high vowels were the stimulus, (lower) mid vowels were more often given as response (4 times) than low vowels (2 times). A chi-squared test performed in R (R Core Team, 2015) shows that this pattern is not random (χ^2 (25) = 6075.6, $p < 0.001$).

As the participants of the production experiment (see chapter 2.2) and the participants of the preceding artificial language learning experiment were all Northern German native speakers, I performed the analysis of the perception experiment again on a subset of the data. This subset contained the data of all 21 participants who had been grown up in Northern Germany (most of them in North Rhine-Westphalia, some of them in Lower Saxony, Berlin, Saxony-Anhalt or Brandenburg). The results are similar to the results of all German native speakers: The percentages of correct responses to high vowels ([i]: 93%, [ĩ]: 85%) was significantly higher than the percentages of correct responses to mid ([ɛ]: 91%, [ɛ̃]: 71%) (Est. = −0.8389, SE = 0.2341, z = −3.584, $p < 0.001$) and to low vowels ([a] = 98%, [ã]: 58%) (Est. = −0.9640, SE = 0.2316, z = −4.162, $p < 0.001$). There was no significant difference between the percentages of correct responses to mid and low vowels (Est. = 0.1251, SE = 0.2042, z = 0.612, p = 0.540).

2.3.3 Discussion

The results confirm my expectations that German native speakers are more likely to confuse [a] with [ã] and [ɛ] with [ɛ̃] than [i] with [ĩ]: The perception of high vowel nasalization is easier than the perception of non-high vowel nasalization.

This pattern is in agreement with the perceptual similarity between non-high oral and nasal vowels, and the dissimilarity between high oral and nasal vowels as shown by Bond (1976) and House & Stevens (1956) – even in a language without phonemic vowel nasalization. These differences can also be observed in the vowel chart presented in Figure 7 as oral and nasal high vowels are more distant from each other than oral and nasal mid or low vowels. In accordance with the production experiment (see chapter 2.2) (Northern) German native speakers are thus guided by phonetics when perceiving unfamiliar sound patterns as for example nasal vowels of different height. This is in line with the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006) and gives evidence for the ease of perception of high nasal vowels.

Hajek & Maeda (2000) proposed that a preference for low nasal vowels may also be due to perception because low vowels are intrinsically longer which makes the nasalization more salient (Laver, 1994; Lehiste, 1970). Unfortunately, I did not control for the length of the stimuli, as they were sliced out of stimuli recorded for the artificial language learning experiment. However, my data does not confirm Hajek & Maeda (2000)'s claim. I measured the duration of the stimuli and found that in my stimuli the low vowel stimuli were not longer than the mid and high vowel stimuli. Further, I found that the oral vowel stimuli (mean: 264 ms) were generally longer than the nasal vowel stimuli (mean: 159 ms). Moreover, the difference in duration was greatest for [ɛ]-[ɛ̃] (165 ms), followed by [a]-[ã] (107 ms) and [i]-[ĩ] (43 ms). Thus, if the perception of nasalization was based on the vowel duration, the contrast between [ɛ] and [ɛ̃] would have to be perceived best and the contrast between [i] and [ĩ] would have to be perceived worst. This, however, was not the case. The results show exactly the opposite. Thus, the results were not confounded by (intrinsic) duration.

2.4 Artificial language learning experiment

The results of the previous investigations show that the perception and the articulation of vowel nasalization by Northern German native speakers are affected by vowel height. The difference between oral and nasal vowels is perceived better in high vowels than in mid and low vowels. The reason for this is that in general oral and nasal higher vowels are not confused with other as much as oral and nasal lower vowels are, which has also been shown for (Northern) German native speakers (see chapter 2.3). The acoustic analysis, however, showed that the difference between oral and nasal vowels is produced better in mid and low vowels than in high vowels. The reason for this is that in general lower nasal vowels are easier to produce than higher nasal vowels, which has also been shown for Northern German native speakers (see chapter 2.2).

Next I investigated whether the phonetics of nasal vowels also plays a role in phonological rule learning. My aim was to investigate the hypothesis that native speakers of Northern German are more likely to learn a nasalization pattern that is phonetically grounded (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006). I was further interested in whether this phonological learning is more affected by the differences in articulation or by the differences in perception. I additionally investigated whether the generalization behavior during learning may be affected by a further factor: phonetic similarity (Kapatsinski, 2013a).

To test whether perceptual differences, articulatory differences or phonetic similarity affect learning a nasalization pattern, I created an artificial language learning experiment in which the crucial condition was the plural. The participants learned that plural forms ended in a nasal vowel followed by the nasal consonant [m] (see phonological rule in (2)) and that diminutive forms ended in an oral vowel followed by [l]. The diminutives were added in order to conceal the focus of the experiment for the participants and to provide evidence that a vowel is only nasalized in the plural. I chose the context VN because this context was used in the perception experiment (see chapter 2.3) and because in this context vowels are coarticulatory nasalized by Northern German native speakers – although only to a weak degree (see chapter 2.2). VN is hence the context that should induce perceptual ease for high vowels as well as articulatory ease for non-high vowels for (Northern) German native speakers.

(2) $/\mathrm{V}/ \rightarrow [\tilde{\mathrm{V}}]\ /\ \_[\mathrm{m}]$

Questions of biases in learning are usually investigated by means of poverty-of-the-stimulus experiments (Baer-Henney, 2015; Baer-Henney et al., 2015a; Cristià & Seidl, 2008; Finley, 2009, 2012; White, 2013, 2014; White & Sundara, 2014; Wilson, 2003, 2006; Zuraw, 2007). In such experiments participants are exposed to one pattern, but are tested on novel patterns that they had not been exposed to before. I therefore compared three groups of German learners who all learned a nasalization pattern but each with a different vowel height – either with high [i], mid [ɛ] or low [a]. In the test phase the learners were asked to judge plural forms with vowels from all heights.

First, I lay out the predictions based on the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006). If there is a learning bias that is based on ease of perception, I will expect that [i]-learners learn their nasalization pattern best, followed by [ɛ]-learners and [a]-learners and I will expect that [a]-, [ɛ]- and [i]-learners all generalize more to [i]-items than to [ɛ]- or [a]-items. If ease of perception is part of their linguistic knowledge, they will know that high nasal and oral vowels are easiest to distinguish. Having perceived a nasal vowel of any height (and they can do so in principle, see Table 3), they nevertheless have a preference to generalize to high vowels. If there is a learning bias that is based on ease of articulation, I will expect that [a]-learners learn their nasalization pattern best, followed by [ɛ]-learners and [i]-learners and I will expect that [a]-, [ɛ]- and [i]-learners all generalize more to [a]-items than to [ɛ]- or [i]-items.[4] If ease of articulation is part of their linguistic knowledge, they will know that low nasal vowels are easiest to pronounce (Hajek & Maeda, 2000; Ohala, 1975; Zsiga, 2013). Having perceived a nasal vowel of any height, they nevertheless have a preference to generalize to low vowels.

Table 4: Predictions for the artificial language learning experiment testing vowel nasalization.

	Learning	**Generalizations of**		
		[a] -learners	**[ɛ] -learners**	**[i] -learners**
Ease of perception	[i] > [ɛ], [a]	[i] > [ɛ]	[i] > [a]	
Ease of articulation	[a] > [ɛ], [i]		[a] > [i]	[a] > [ɛ]
Phonetic similarity	/	[ɛ] > [i]	[a] > [i]	[ɛ] > [a]

Second, I lay out the predictions based on phonetic similarity (Kapatsinski, 2013a). If there is a learning bias for phonetically similar patterns, I will expect that [a]-learners extend the nasalization more to [ɛ]-items than to [i]-items. This is because [a] resembles [ɛ] more than [i] does (see Figure 7). [ɛ]-Learners

[4] According to the data of the production experiment in chapter 2.2 there should be no difference between [a] and [ɛ]. However, for reasons of simplicity and reliability I base my predictions on the literature, e.g. Bell-Berti (1993) and Ohala (1975), and not on the small subset of German native speakers I tested.

should extend the nasalization to [a]-items more than to [i]-items. This is because [ɛ] is more similar to [a] than to [i] (see Figure 7). [i]-Learners, finally, should extend the nasalization to [ɛ]-items more than to [a]-items. This is because [i] is more similar to [ɛ] than to [a] (see Figure 7). In Table 4 you can find an overview of the different predictions. In this Table empty cells are those in which the vowel in training would be the one most generalized to. In this case the effect of phonetics and the effect of exposure during the training phase of the experiment would not be separable. Cells including '/ ' are those in which the phonetic effect makes no prediction.

2.4.1 Method

Stimuli The artificial language consisted of singular, plural and diminutive forms. The stimuli were constructed from a subset of the German (Wiese, 1996) and Portuguese phoneme inventories (Azevedo, 2005). The items conformed to the phonotactics of German (Wiese, 1996) – with the exception of the nasal vowels. The structure of the stimuli and examples for each grammatical form are illustrated in Table 5.

Table 5: Structure of the stimuli for the artificial language learning experiment testing vowel nasalization.

Form	C_1	V_1	C_2	V_2	Suffix	Example
singular	[p d k ʃ v]	[o u]	[b t g z f]	[i ɛ a]	∅	[koga]
plural	[p d k ʃ v]	[o u]	[b t g z f]	[ĩ ɛ̃ ã]	[m]	[kogãm]
diminutive	[p d k ʃ v]	[o u]	[b t g z f]	[i ɛ a]	[l]	[kogal]

The singulars were $C_1V_1C_2V_2$-items with a high, mid or low vowel as the second vowel (V_2). To form a plural, an [m] was suffixed, which caused the nasalization of the preceding vowel, and to form a diminutive an [l] was suffixed without any other phonological change. The consonants [p], [d], [k], [ʃ] and [v] were used in C_1-position and the consonants [b], [t], [g], [z] and [f] were used in C_2-position. V_1 was one of the back vowels [o] and [u] whereas V_2 was one of the front vowels [i], [ɛ] and [a]. The oral vowels [i], [ɛ] and [a] and their nasal counterparts [ĩ], [ɛ̃] and [ã] were used as the critical vowels because these vowels are part of the Portuguese phoneme inventory (Azevedo, 2005) and they allowed me to compare vowels of different heights.

For the training phase I used 48 items for each group of participants (16 singulars, 16 plurals and 16 diminutives). For the subsequent test phase I used 48 stimulus pairs, each consisting of a form which conformed to the nasalization rule (e.g. the correct plural [kogãm] and the correct diminutive [dufil]) and of a form which did not conform to the nasalization rule (e.g. the incorrect plural [kogam] and the incorrect diminutive [dufĩl]). Half of the pairs (n = 24) tested the plural formation and half of the pairs (n = 24) tested the diminutive formation. In both stimulus groups there was an equal number of eight pairs for each of the three vowel heights. Half of these eight pairs were part of the training items (n = 4) and half of them were not part of the training items (n = 4). A complete list of all stimuli used in this experiment can be found in the appendix (see Table 27-30).

The stimuli were recorded by a female fully bilingual native speaker of Portuguese and German who also did the recording for the perceptual confusion experiment. Recording took place in an anechoic booth in the phonetics laboratory of the Heinrich-Heine University Düsseldorf to avoid disturbing noise. For the recordings a microphone (phantom power 48V), a Sound Devices amplifier, a Marantz recorder and a Transtec computer were used. The sampling rate was 48 kHz. Stimuli were recorded as the answer to the Portuguese question *O que é que eu disse?* 'What did I say?' to focus on the stimulus item and to ensure a uniform language environment that allows the speaker to naturally produce nasal vowels. The target sentence itself was not read aloud but quietly. After recording all stimuli were sliced out at the nearest zero-crossing point using Praat (Boersma & Weenink, 2019) to avoid irritating noise at the beginning or at the end of the items. The intensity of the stimuli was adjusted to 70 dB using Praat (Boersma & Weenink, 2019).

Procedure The artificial language learning experiment was divided into a perceptual training phase and a perceptual forced-choice test. The experiment was run using the software PsychoPy (Peirce, 2007) on a Windows laptop. Participants listened to the auditory stimuli via headphones. The experiment lasted 10 to 15 minutes and took place in an anechoic booth in the phonetics laboratory of the Heinrich-Heine University Düsseldorf to avoid disturbing noise.

In the experiment the poverty-of-the-stimulus method (Wilson, 2006) was used, which allowed me to study not only the learning behavior but also the generalization behavior of the participants. The participants were trained on a subset of the stimuli but tested on all stimuli – including stimuli known from training and stimuli they had not yet heard. There were three experimental groups: one for each vowel height of the stimuli's V_2. During training each group heard the nasalization of only one vowel – either [a] or [ɛ] or [i]. The test was identical for each group and included high, mid and low vowels.

At the beginning of the experiment participants were informed that a new language was discovered on an island in the South Sea and that they would get to know this new language as well as the animals living on this island in the course of the experiment.

In a short introductory phase the participants were familiarized with the set-up of the experiment. They listened to three German animal names in the singular, plural and diminutive, e.g. *Hase* 'rabbit', *Hasen* 'rabbits', *Häschen* 'little rabbit'. These examples were supported visually by means of images with these animals. The images I used are part of the Snodgrass & Vanderwart (1980) collection and illustrate the meaning of the forms in the artificial language.

After the introductory phase the training and test phase followed. The participants were told that the experiment consists of two parts, a *first phase* and a *second phase* (similar to Wilson (2006)). During training the participants heard two repetitions of 48 stimuli (16 singulars, 16 plurals and 16 diminutives) in randomized order, which should mimic word learning (van de Vijver & Baer-Henney, 2014). An auditory stimulus was played while the visual stimulus was shown for 1000 ms. As I used an artificial language I showed fantasy animals as visual stimuli (van de Vijver & Baer-Henney, 2014). Therefore a singular form was accompanied by an image of a single fantasy animal, a plural form was accompanied by an image of two fantasy animals and a diminutive form was accompanied by an image of a small fantasy animal. There was an inter-stimulus-interval of 500 ms (see Figure 12). During training participants were confronted with positive input only, which means that they never listened to an incorrect plural, e.g. [kogam], or to an incorrect diminutive, e.g. [dufĩl].

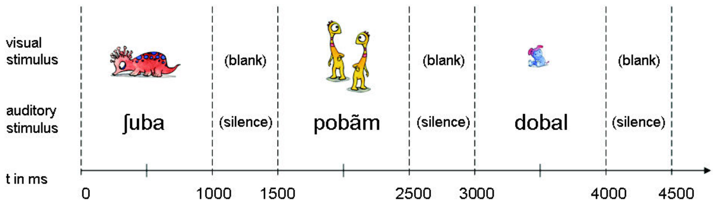

Figure 12: Time course of the training phase in the artificial language learning experiment.

The subsequent test was a forced-choice task and was identical for all groups. There were 48 stimulus pairs that consisted of a correct and an incorrect form that only differed in the nasalization of V_2. Half of the pairs (n = 24) tested the plural formation and half of the pairs (n = 24) tested the diminutive formation. There was for instance a correctly nasalized plural form like [kogãm] vs. an incorrectly oral plural form like [kogam] and an incorrectly nasalized diminutive form like [dufĩl] vs. a correctly oral diminutive form like [dufil]. There were 16 pairs with high, mid and low vowels each. A trial consisted of the presentation of the first form, which lasted one second, followed by the presentation of the second form (see Figure 13). There was an equal number of trials in which the first and the second form was the correct one. The inter-stimulus-interval was 200 ms. During the auditory presentation the corresponding visual support was displayed. After that the participants had 3000 ms to decide which of the two forms was correct by pressing either the right or the left arrow key. After an inter-trial-interval of 500 ms the next stimulus pair was presented. The stimulus pairs were presented in randomized order.

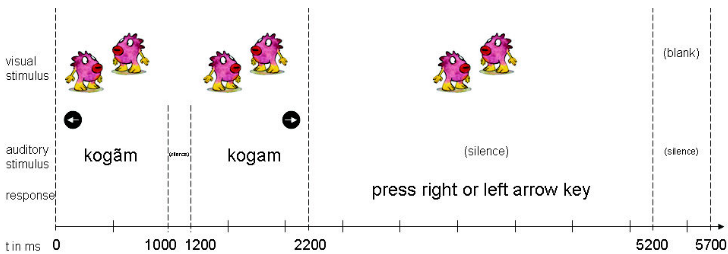

Figure 13: Time course of one trial in the test phase of the artificial language learning experiment.

Participants 61 native speakers of Northern German took part in the experiment (39 identified as female, and 22 identified as male, mean age: 28.0, range: 18-74). They were randomly assigned to one of the three experimental groups. 20 participants were trained with nasalization of the high vowel [i], 20 participants were trained with nasalization of the mid vowel [ɛ] and 21 participants were trained with nasalization of the low vowel [a]. All groups were tested with all vowels [i], [ɛ] and [a]. No one reported knowledge of a language that uses nasalized vowels distinctively. All participants had normal or corrected-to-normal vision and no hearing problems. None of them had participated in the previous experiments (perception experiment and production experiment). Participants were given a small expense allowance for their participation. Prior to the experiment each participant filled out a consent form and a questionnaire with which data about the participants' age, sex, native language, foreign language skills, the region they had been grown up and education was collected.

2.4.2 Results

From 2928 data points (61 participants × 48 trials) 51 were not analyzable because the participants did not respond within 3000 ms. I analyzed the remaining 2877 data points with R (R Core Team, 2015) by means of a generalized linear mixed-effects analysis with the corresponding package lme4 (Bates et al., 2015). I was interested in the learning behavior and in the generalization behavior across experimental groups. For an overview of the results to plural forms only (n = 1442) see Figure 14[5]. This Figure was created using R(R Core Team, 2015) and the R packages Rmisc (Hope, 2013) and ggplot2 (Wickham, 2009).

First of all, I compared the results of the trained vowel in the plural items in each group. This is what I call *learning*. In the model that fitted the data best (as assessed by backward step-wise elimination (Baayen, 2008)) ACCURACY (correct and incorrect) served as dependent variable and TRAINED VOWEL ([a], [ɛ] and [i]) was the independent variable. PARTICIPANT and ITEM were included as random intercepts (R syntax: glmer(ACCURACY

[5]I decided to focus on plural forms only as plural was the crucial condition. Besides, the results including plurals and diminutives showed the same tendencies as plural only results. However, the differences were not that robust so that some of them did not reach significance: ([a]-learners: [a] (63%) = [ɛ] (62%) = [i] (59%), [ɛ]-learners: [ɛ] (84%) > [a] (70%) = [i] (68%), [i]-learners: [i] (84%) > [ɛ] (71%) > [a] (56%)

$\sim$ TRAINED VOWEL + (1|PARTICIPANT) + (1|ITEM), family = "binomial"))

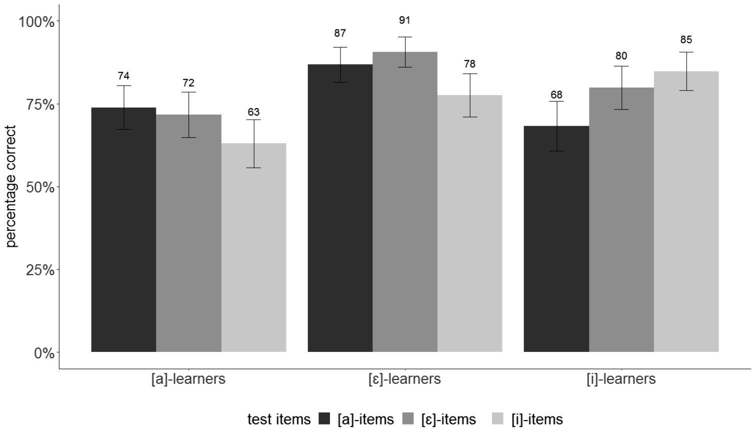

Figure 14: Percentages of correct responses ± 1.96 SE to plural items across all training groups and all vowel heights at test.

In the trained condition the percentage of correct responses of [ɛ]-learners was 91%, that of [i]-learners was 85% and that of [a]-learners was 74%. The percentages of correct responses of [ɛ]- and [i]-learners did not differ significantly from each other (Est. = −0.6072, SE = 0.4929, z = −1.232, p = 0.218). [ɛ]-learners gave significantly more correct responses to [ɛ]-items than [a]-learners to [a]-items (Est. = −1.3880, SE = 0.4662, z = −2.977, p < 0.01) and [i]-learners gave marginally significantly more correct responses to [i]-items than [a]-learners to [a]-items (Est. = −0.7808, SE = 0.4377, z = −1.784, p = 0.075). The percentage of correct responses to the trained items in the plural differed from chance level for [ɛ]-learners (Est. = 2.5953, SE = 0.3938, z = 6.591, p < 0.001), for [i]-learners (Est. = 1.9881, SE = 0.3541, z = 5.615, p < 0.001) and for [a]-learners (Est. $= 1.2073$, SE $= 0.2898$, $z = 4.166$, $p < 0.001$).

I also investigated how participants judged items with untrained vowels. This is what I call *generalization*. Therefore I performed three generalized linear mixed-effects models – one for each learning group. In the models that fitted the data best (as assessed by backward stepwise elimination (Baayen, 2008)) ACCURACY (correct and incorrect) served as dependent variable and VOWEL AT TEST ([a], [ɛ] and [i]) as independent variable. PARTICIPANT and ITEM were included as random intercepts (R syntax: glmer(ACCURACY ~ VOWEL AT TEST + (1|PARTICIPANT) + (1|ITEM), family = "binomial"))

The first analysis is done on the data of the [a]-learners. Participants provided most correct answers to [a]-items (74%). They provided slightly less correct answers to [ɛ]-items (72%), but not significantly so (Est. $= -0.1298$, SE $= 0.2606$, $z = -0.498$, $p = 0.619$). They provided significantly less correct answers to [i]-items (63%) than to [a]-items (Est. $= -0.6255$, SE $= 0.2563$, $z = -2.440$, $p < 0.05$) and to [ɛ]-items (Est. $= -0.4957$, SE $= 0.2528$, $z = -1.961$, $p < 0.05$). The percentage of correct responses to [i]-items (Est. $= 0.6724$, SE $= 0.3037$, $z = 2.215$, $p < 0.05$) and to [ɛ]-items (Est. $= 1.1682$, SE $= 0.3113$, $z = 3.753$, $p < 0.001$) both differed significantly from chance level.

The second analysis is done on the data of the [ɛ]-learners. Participants provided most correct answers to [ɛ]-items (91%). They provided slightly less correct answers to [a]-items (87%), but not significantly so (Est. $= -0.4343$, SE $= 0.4413$, $z = -0.984$, $p = 0.325$). They provided significantly less correct answers to [i]-items (78%) than to [ɛ]-items (Est. $= -1.2082$, SE $= 0.4214$, $z = -2.867$, $p < 0.01$). The [ɛ]-learners provided only marginally significantly more correct answers to [a]-items than to [i]-items (Est. $= -0.7739$, SE $= 0.3965$, $z = -1.952$, $p = 0.051$). The percentage of correct responses to [a]-items (Est. $= 2.3908$, SE $= 0.4267$, $z = 5.603$, $p < 0.001$) and to [i]-items (Est. $= 1.6169$, SE $= 0.3915$, $z = 4.130$, $p < 0.001$)

differed significantly from chance level.

The third analysis is done on the data of the [i]-learners. Participants provided most correct answers to [i]-items (85%). They provided slightly less correct answers to [ɛ]-items (80%), but not significantly so (Est. $= -0.3615$, SE $= 0.4065$, z $= -0.889$, p $= 0.374$). They provided significantly less correct answers to [a]-items (68%) than to [i]-items (Est. $= -1.0496$, SE $= 0.3951$, z $= -2.657$, p < 0.01). [i]-Learners provided only marginally significantly more correct answers to [ɛ]-items than to [a]-items (Est. $= -0.6881$, SE $= 0.3812$, z $= -1.805$, p $= 0.071$). The percentage of correct responses to [a]-items (Est. $= 0.8410$, SE $= 0.2838$, z $= 2.964$, p < 0.01) and to [ɛ]-items (Est. $= 1.5291$, SE $= 0.3079$, z $= 4.965$, p < 0.001) differed significantly from chance level.

2.4.3 Discussion

The artificial language learning experiment shows that native speakers of Northern German are able to learn a vowel nasalization pattern. Participants performed not only in their trained vowel condition (learning) but also in their untrained vowel conditions (generalizations) above chance level. There is, nevertheless, a clear effect of training as all participants performed best in their trained condition.

Their performance in the trained conditions differed regarding vowel height: Participants trained with high [i] and with mid [ɛ] learned their vowel nasalization pattern better than participants trained with low [a]. This result is partly in line with the predictions based on ease of perception. Their performance in the generalizations differed regarding vowel height: Participants trained with low [a] generalized their learned vowel nasalization pattern more to mid [ɛ] than to high [i]. This result is in line with the predictions based on phonetic similarity. Participants trained with mid [ɛ] generalized their learned vowel nasalization pattern more to low [a] than to high [i]. This result is line with the predictions based on phonetic similarity and with the predictions based on articulatory ease. Participants trained with high [i] generalized their learned vowel nasalization pattern more to mid [ɛ] than to low [a]. This result is in line with the predictions based on phonetic similarity.

Comparing the results of the learning experiment with the predictions, I can say that the hypothesis that (morpho)phonological learning is affected by ease of articulation is not confirmed[6] and that the hypothesis that morphophonological learning is affected by ease of perception is partly confirmed. The hypothesis based on ease of articulation predicts a learning advantage for [a]-learners over [ɛ]- and [i]-learners. I, however, found the opposite. The hypothesis based on ease of perception predicts a learning advantage for [i]-learners over [ɛ]- and [a]-learners. However, I found a learning advantage for [i]- and [ɛ]-learners over [a]-learners. I can further say that the hypothesis that morphophonological generalizations are affected by ease of perception is never confirmed and that the hypothesis that morphophonological generalizations are affected by ease of articulation are confirmed for [ɛ]-learners, but not for [i]-learners.[7] However, the best fit is between the predictions based on phonetic similarity and the results of the experiment: All of its predictions are confirmed. In Table 6 an overview of which predictions are confirmed by the experimental results and which are not can be found. Green check marks represent confirmed predictions and red x marks represent unconfirmed predictions. Empty cells are those in which the vowel in training would be the one most generalized to. In this case the effect of phonetics and the effect of exposure during the training phase of the experiment would not be sepa-

[6]Based on the data of the production experiment in chapter 2.2 [a] and [ɛ] do not differ in ease of articulation. The predictions for learning based on ease of articulation would then be [a], [ɛ] > [i], which is not confirmed.

[7]Based on the data of the production experiment in chapter 2.2 [a] and [ɛ] do not differ in ease of articulation. The predictions for generalizations based on ease of articulation would then be [ɛ] > [i] for [a]-learners, [a] > [i] for [ɛ]-learners, [a] = [ɛ] for [i]-learners, which means that the predictions for the [a]-learners and for the [ɛ]-learners are confirmed but not the predictions for the [i]-learners.

rable. Cells including '/ ' are those in which the phonetic effect makes no prediction.

I therefore conclude that the results are best explained by assuming that learners are best at learning the perceptually easy pattern and that learners generalize from observed patterns to phonetically similar patterns, but that my findings provide only limited evidence for a learning bias as a result of phonetic markedness.

Table 6: Predictions for the artificial language learning experiment testing vowel nasalization and their evaluation.

	Learning	Generalizations of		
		[a] -learners	**[ɛ] -learners**	**[i] -learners**
Ease of perception	[i] > [ɛ], [a] (✓)	[i] > [ɛ] X	[i] > [a] X	
Ease of articulation	[a] > [ɛ], [i] X		[a] > [i] ✓	[a] > [ɛ] X
Phonetic similarity	/	[ɛ] > [i] ✓	[a] > [i] ✓	[ɛ] > [a] ✓

2.5 General discussion

I set out to investigate whether the phonetic factors ease of perception or ease of articulation affect morphophonological learning and generalizations, as predicted within the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006), or whether the phonetic factor similarity affects morphophonological generalizations. The first hypothesis claims that there is a preference for learning phonetically grounded phonological alternations over phonetically arbitrary alternations (Wilson, 2006). This preference in learning is revealed cross-linguistically: Typologically common patterns are those that are preferred by the learner over typologically rare patterns. The second hypothesis claims that learners generalize from observed alternations in learning to phonetically similar alternations in novel words (Kapatsinski, 2013a).

In preceding studies I investigated what these hypotheses mean for vowel nasalization patterns of different height. First, I checked whether it is true for vowel nasalization patterns that the phonetically marked patterns are the typologically rare patterns. The results of the typological study showed that vowel nasalization is equally distributed across vowel heights: About 20% of all vowels (independent of whether they are high, mid or low) are nasal vowels. Thus, the preference for articulatory and perceptually easy patterns is not revealed typologically – speaking against the hypothesis of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006). As a consequence, a phonetic bias does not necessarily have to be encoded as markedness in our grammar. Second, I tested which of the vowels in our stimuli are phonetically most similar to each other. The results of the acoustic analysis revealed that low [a] and lower mid [ɛ] are more similar to each other than lower mid [ɛ] and high [i]. This is true for the oral vowel pairs as well as for the nasal vowel pairs. Third, I tested whether Northern German native speakers, who are not familiar with vowel nasalization, are affected by ease of articulation. The production experiment confirmed that low and mid nasal vowels are more easily articulated than high nasal vowels. Fourth, I tested whether (Northern) German native speakers, who are not familiar with vowel nasalization, are affected by ease of perception at all. The perception experiment confirmed that high nasal vowels are more easily distinguished from high oral vowels than mid or low nasal vowels from mid or low oral vowels. Thus, not only Northern German native speakers' articulation and perception may be affected by ease of perception and ease of articulation but also Northern German native speakers' learning and generalization behavior.

I conducted an artificial language learning experiment in which three groups of Northern German learners were taught a language in which plural was expressed by [ãm] ([a]-learners) by [ɛ̃m] ([ɛ]-learners) or by [ĩm] ([i]-learners). I subsequently tested the participants on their generalizations to plurals with all vowels.

Regarding learning the results support the role of ease of perception. Participants who were exposed to the perceptually easy pattern learned their trained pattern best, whereas participants who were exposed to the articulatory easy pattern learned their trained pattern worst. This is in line with the results of Glewwe (2019) who found that only a perceptual bias and not an articulatory bias influences phonological learning. Regarding generalizations the results confirm the predictions based on phonetic similarity for all learning groups, whereas the predictions based on ease of perception were never confirmed and the predictions based on ease of articulation were only confirmed for [ɛ]-learners.

Thus, the results suggest a grammar in which perception biases learning and in which generalizations are formed on the basis of phonetic similarity. When participants were trained on plurals with nasal vowels, all of them perceived the nasalization in the plurals. However, participants trained on plurals with nasal [i] or with nasal [ɛ] perceived the nasalization better than participants trained on plurals with nasal [a], as oral and nasal [a] are perceptually more similar to each other than oral and nasal [i] or [ɛ]. Thus, [i]- and [ɛ]-learners were more sure about the phonological alternation they had learned than [a]-learners were. When the participants were trained with plurals with the vowel [a] they took those plurals as good examples of a plural (Booij, 2010; Croft, 2007; Jackendoff & Audring, 2016; Kapatsinski, 2013b). In the test phase they would obviously accept those plurals as good instances of a plural form. Plurals with an [ɛ] are less good examples of plurals, but better examples than plurals with an [i] because an [a] resembles an [ɛ] more than an [i] (Kapatsinski, 2013a) (see Table 2 and Figure

7). A similar logic explains the results of the [ɛ]-learners. For them a plural with an [a] is a better example of a plural than a plural with an [i] because an [a] resembles an [ɛ] more than an [i]. For [i]-learners a plural with an [ɛ] is a better example of a plural than a plural with an [a] because an [ɛ] resembles an [i] more than an [a].

The results document an effect of phonetics on phonology and not a purely phonological similarity effect. Even though one could analyze [a] and [ɛ] as [−high] and [i] as [+high], the results are not simply effects of [−high] vowels patterning together in contrast to [+high] vowels. [i]-Learners generalized more to [ɛ] than to [a]. If I had been dealing with a phonological effect, I would have expected that [ɛ] and [a] would be treated in the same way. There are other ways of phonologically classifying the vowels in the experiment, e.g. grouping [ɛ] and [i] together as [−back] vowels, but none of these classifications explains the results of the experiment as well as plain phonetic similarity.

The results of the learners do not mirror the typology of nasal vowels and it is even not the case that the phonetic factors ease of articulation and ease of perception for nasal vowels (Hajek & Maeda, 2000; Ohala, 1975; Zsiga, 2013) match the typological distribution of nasal vowels when considering the percentages of nasal vowels on all vowels (see Table 1). However, when considering the raw numbers of nasal vowels – independent of the number of oral vowels – the phonetic factors perfectly match the typological distribution (see Table 1). In my opinion, however, this latter account with raw numbers is misleading and wrong. Thus, ease of articulation and ease of perception are not represented in phonology as markedness constraints – in the case of vowel nasalization. However, it may be the case that the nasalization data I used from the UPSID corpus (Maddieson, 1984) is not representative as it contains only phonemic nasal vowels and not allophonically nasalized vowels. Considering allophonically nasalized vowels as well, it might be the case that one can observe an effect of ease of perception and ease of articulation even in the percentages. To investigate this, however, is beyond the scope of this dissertation. The results of the artificial language learning experiment therefore show – independent of whether percentages or raw numbers are considered – that learning and especially generalizing morphophonological patterns is not necessarily guided by markedness. It actually may be a tall order for the results of an artificial language learning experiment which takes about 15 minutes to complete to reflect typological distributions (Blevins, 2004; Moreton & Pater, 2012b; Wilson, 2006) but it is still possible that ease of perception and ease of articulation affect languages diachronically (Blevins, 2004). Sound patterns in languages are the result of eons of change that is affected by many factors only one of which is the phonetic grounding of a sound pattern. Therefore, not all learning biases directly affect typological patterns.

It is important to stress, though, that there are phonetic factors that influence synchronic morphophonological generalizations. The learners generalized from observed alternations to novel words, in a way that reflects the phonetic similarity between the observed patterns and the new patterns. This result is important because it provides data that undermines the hypothesis that phonology is free of phonetic considerations (Blevins, 2004; Hale & Reiss, 2000; Ohala, 1986). In the most recent instantiation of this hypothesis Blevins (2004) maintains that the role of phonetic explanation in phonology is restricted to the diachronic domain. My results and the results of Wilson (2006) show that learners do use phonetics to make phonological synchronic generalizations (White, 2017; Zuraw, 2007). The proposal that all phonetic explanation in phonology is diachronic is too strong, and learners clearly use phonetic similarity to make phonological generalizations.

I thus argue that phonetics needs to be more tightly integrated into phonology as my results support the hypothesis that phonological representations must include phonetic details (Flemming, 2001). According to Flemming (2001) allophonic nasalization is often described as phonetic because it is automatically achieved by lowering the velum during the articulation of a vowel in the context of a nasal consonant. However, the same process is found in phonology when the contrast between oral and nasal vowels is neutralized in the context of nasal consonants in many languages. Thus, phonetics and phonology should not be regarded as two independent components of our grammar. They are rather two components that largely overlap and whose boundary is arbitrary. This can be illustrated with the results of my experiments. Phonetically high oral and nasal vowels are more different from each other than low oral and nasal vowels. Due to these small phonetic differences among vowel heights the phonological process of vowel nasalization is more easily learned for high vowels than for low vowels. The phonetic difference is hence part of the phonological representations and affect how they are learned. They are further used during the generalization of one phonological pattern to another one. Interestingly, phonetic details affecting learning and generalizations can be different, with one phonetic detail, e.g. perceptual ease, affecting learning and another phonetic detail, e.g. phonetic similarities, affecting generalizations. To conclude, in our grammar phonetic details, such as phonetic similarities and perceptual ease, are stored which are used during learning phonological alternations. This enabled the learners to use phonetic details in generalizing a pattern to novel items as well. Without overlapping modules of phonetics and phonology the experimental results could not be explained properly.

2.6 Conclusion

I started out by asking whether learners take phonetic similarity between sounds of phonological patterns into account when generalizing to novel patterns with other sounds. I found that they do, and that their results shed light on the role of phonetics in morphophonological learning. Learners generalize from observed morphophonological patterns to phonetically similar patterns. This hypothesis explains my results much better than the hypothesis that ease of perception or ease of articulation affect morphophonological generalizations.

However, during learning the participants seem to have been biased by perceptual ease. In this way, the results show that phonetic details are used in learning about morphophonological alternations as well. Even though I found no link between generalizations and typology, it is not the case that phonetic similarity is irrelevant to learning synchronic morphophonological patterns.

3 Articulation and learning of vowel nasalization

3.1 Introduction

The previous artificial language learning experiment tested only the perception of nasalization. However, I expect that the perception of nasalization also alters the production of nasalization (Zellou et al., 2016, 2017). It is well known that speakers accommodate their speaking style to their interlocutors (Pardo, 2006; Pardo, Jay & Krauss, 2010), which means that the way people speak is perceived by the hearers, who, in turn, modify their own speech by imitating their partners (Sancier & Fowler, 1997).

This effect is also present in the laboratory setting in so-called shadowing tasks (Goldinger, 1998). Participants read aloud a word (target), then listen to another word (prime) and then read aloud the target again. In those studies the effect of imitation is studied by AXB perception experiments in which listeners should judge whether the target read aloud before hearing the prime or the target read aloud after hearing the prime is a better fit to the prime (X). It was found that the target read aloud after the prime was a better fit to the prime than the target read aloud before hearing the prime.

Nielsen (2011) found that the imitation of a prime is also generalized to another consonant and thus showed that people generalize the coarticulatory properties of an auditorily presented word to other, similar words, which have not been auditorily presented. Her study consisted of an auditory exposure phase and two reading session – one before and one after exposure. In the auditory exposure phase participants listened to items starting with /p/ whose VOTs were manipulated; it was increased. The results showed that after exposure participants produced longer VOTs in items starting with /p/ – familiar items with /p/ and novel items with /p/ – and in novel items starting with /k/. Thus, spontaneous phonetic imitation can be generalized to novel words and to novel phonemes.

Similar experiments have also been conducted by testing vowel nasalization. Zellou et al. (2017) tested how the degree of nasalization in the auditory prime affects the pronunciation of nasal vowels. During the experiment native speakers of American English listened either to a naturally nasalized prime or to a hypernasalized prime and were then asked to read aloud an existing English word with a VN-sequence displayed on a computer screen. Hypernasalized primes were created by replacing the vowel in the naturally nasalized CVN-sequence by a combination of this vowel with a vowel from an NVN-sequence. The results showed that the increase in produced nasality – measured by means of $A1-P0$ – was greater after auditory exposure to hypernasalized primes than after exposure to naturally nasalized primes. Interestingly, when listening only to naturally nasalized primes, speakers did hardly increase their produced nasality. However, when listening to naturally nasalized primes and to hypernasalized primes, speakers increased their produced nasality after exposure. Conversely, speakers listening to hypernasalized primes and naturally nasalized primes produced less nasality than speakers listening only to hypernasalized primes after exposure. This suggests that the imitation of nasalization is influenced by overall experience and not only by the recent prime. Zellou et al. (2017) interpreted it in the light of Goldinger (1998)'s theory who claimed that low-frequency prime words (= less common words, those with hypernasality) were better imitated than high-frequency ones (= common words, those with natural nasality) (Goldinger & Azuma, 2004). Similar results have been found by Zellou et al. (2016): Participants produced the vowels after exposure to hypernasalized auditory primes with greater nasality than before exposure.

Goldinger (1998) explained this in the light of an episodic theory (Hintzman, 1984, 1988), also known as exemplar theory. In this theory it is assumed that each word we hear leaves a memory trace that also includes phonetic details. Whenever we hear a word, several stored traces are activated depending on the phonetic similarity to the heard word. As a consequence we interpret it in the processes of word recognition as the word, whose traces are activated most. When it comes to the production of this word or a similar word right after the auditory presentation of the prime, those traces are still activated and thus influence the production as well. As low-frequency words are represented by fewer traces, the influence of a low-frequency prime on the activation of a single representation is greater – due to missing concurrent words – than the one of a high-frequency prime, and this induces a better imitation of this prime in production. This does not only work for the perception of the whole word, but also for phonetic details, e.g. coarticulatory properties, in this word: As Zellou et al. (2017) concluded, listeners generalize phonetic details of a given word to other – similar – words (Nielsen, 2011). Thus, the perceived hypernasal articulation of the vowels in their experiment was more salient than the naturally nasalized articulation for American English listeners. Consequently, the listeners imitated the uncommon hypernasalization better than the natural nasalization.

Based on these results, there might also be goodness-differences in the imitation of vowel nasalization depending on whether listeners are exposed to high, mid or low vowel nasalization. As German native speakers perceive (strongly) nasalized vowels most often in French loan words, and as there are no high nasal vowels in French (Fagyal et al., 2006), I expect high nasal vowels to be low-frequency primes and mid and low nasal vowels to be high-frequency primes.

I conducted an experiment to test how auditory exposure to nasalization of different vowel heights changes the degree of nasalization in the speech of the hearers. To this end, I exposed participants to auditorily presented vowel nasalization and measured an improvement in production by two reading sessions – one before exposure and one after exposure.

The exposure phase was exactly the same as in the previous artificial language learning experiment (see chapter 2.4), which means that there were three different exposure groups. One group listened to stimuli with high nasal vowels, one group listened to stimuli with mid nasal vowels and one group listened to stimuli with low nasal vowels. The reading tasks were identical for all groups and consisted of sentences containing French loan words. This was done to elicit vowel nasalization, since I assumed that vowel nasalization is produced stronger in French loan words than in a nasal context in a word of German origin – remember that coarticulatory vowel nasalization is weak in German (see the data of the production experiment on vowel nasalization in chapter 2.2). Moreover, French loan words can be adapted to the German language to different degrees by means of vowel quality shifts (in the presence or absence of a nasal consonant), which might give further hints to the degree of nasalization. More details about these quality shifts are provided below.

The previous artificial language learning experiment (see chapter 2.4) consisted of a perceptual training phase and a perceptual test phase. Thus, it tested perceptual effects of learning vowel nasalization only. I was also interested in whether similar effects can be found in articulation. As there is no phonemic vowel nasalization in German (Wiese, 1996), German native speakers are appropriate subjects to test this. However, I was faced with the difficulty of training someone an articulation she does not know. For reasons of simplicity I decided to keep the perceptual training phase – which is called *exposure phase* in this chapter – but to acoustically measure an improvement in production in recordings from two reading sessions – one before exposure and one after exposure.

As there are no phonemic nasal vowels in German – especially not in the varieties spoken in the Northern part of Germany –, German native speakers have to imitate nasalization when articulating French loan words (Laeufer, 2010; Wiese, 1996). The difference in the $A1-P0$ value as used in the acoustic analysis described in chapter 2.2 is thus not the only measurement for vowel nasalization by German native speakers. Laeufer (2010) investigated the pronunciation of French loan words by German native speakers and named three types of imitations which differ in the degree of similarity to a true nasal vowel: pronouncing a single (strongly or weakly nasalized) nasal vowel without a following nasal consonant [Ṽ], pronouncing an oral vowel followed by a velar nasal consonant [Vŋ] and pronouncing an oral vowel followed by an alveolar or bilabial nasal consonant [VN]. [Ṽ] is considered to be the most equal equivalent to the French nasal vowel, although the nasalization produced by German native speakers can be weaker than the one produced by French native speakers. [Vŋ] is considered to be a better imitation of French vowel nasalization than [VN] because [ŋ] is more similar to a nasal vowel than [n] or [m] in terms of articulation, acoustics and perception (House, 1957). Greisbach (2003) observed in his study, for which he recorded French loan words by German native speakers, one additional option: a (strongly or weakly) nasalized vowel followed by a nasal consonant [ṼN]. He further added that the quality of the vowel can shift. According to Laeufer (2010) the degree of imitation depends on various factors, e.g. education, geographical origin, register, everyday use and frequency. The higher the education, the more south someone has grown up, the more formal the situation, the less often the word is used and the less frequent it is, the more precise is the imitation of nasal vowels and thus the greater is the degree of nasalization. Greisbach (2003) observed an influence of education as well: The more educated the speakers are, the more often they pronounce nasalized vowels. He further added that vowel quality and its position in the words do not explain the different degrees of nasalization, and that German native speakers do not seem to use rules for pronouncing the vowels in French loan words. It is rather very variable. In this study I controlled for the factors education, geographical origin and register, whereas the factors everyday use and frequency may have confounded the results (for details see chapters 3.2 and 3.2.2).

Greisbach (2003) and Laeufer (2010) further observed pronunciation variants with different vowels depending on the degree of nasalization. This is in accordance with my own impression and can be explained with the acoustics and the perception of nasal vowels. Due to the nasalization F1 is lowered in low vowels, but raised in non-low vowels (House, 1957), which leads to a centralized vowel space for nasal vowels compared to the vowel space of oral vowels. As F1 correlates with tongue height, this acoustic modification can change the perception of height in nasal vowels. Experimental studies (Beddor, 1993; Beddor, Krakow & Goldstein, 1986; Fant, 1960; Hecker, 1962; Wright, 1975) showed that low nasal vowels are perceived as higher than their oral counterparts and that high nasal vowels are perceived as lower than their oral counterparts. Carignan (2018) used ultrasound to test whether this acoustic modification is directly linked to the articulation and found that low and mid nasal vowels are articulated with a more raised and retracted tongue body than their oral counterparts, whereas this is not true for the high vowel /i/. There is also evidence from English (Carignan, Shosted, Shih & Rong, 2011; Pruthi & Epsy-Wilson, 2004) and Hindi (Shosted, Carignan & Rong, 2012) showing that the position of the tongue in oral vowels and their nasal counterparts is not the same (Styler, 2015). Therefore, I suggest that German native speakers make use of a further type of imitation for nasal vowels: the shift of vowel quality. The lower high vowels are produced and the higher low vowels are produced, the better the imitation of vowel nasalization. To measure the degree of nasalization I thus used different means: the choice of vowel, the shift of formants (F1 (and F2)) and differences in A1−P0 values.

Based on the phonetics of nasalization, the different theoretical considerations (see also the predictions for the previous artificial language learning experiment in Table 4) and the results of the previous artificial language learning experiment (see chapter 2.4) I made the following predictions for the outcome of this production study. First of all, I predicted that there is no difference across groups before exposure. All groups show the strongest nasalization in low vowels due to articulatory ease. This would be in line with the results of the production experiment in chapter 2.2. Second, I predicted that there is an increase of nasalization after exposure. Based on the results of the previous artificial language learning experiment (see chapter 2.4.2) I predicted that the participants in the present experiment will learn the nasalization they will be exposed to as well because the training phase of the previous artificial language learning experiment and the exposure phase of the present experiment were identical. Consequently, I predicted that the participants increase the nasalization of their exposure vowel. This increase might be affected by perceptual frequency. To briefly recapitulate (for details see chapter 1.3): Perceptual frequency refers to how often we are exposed to a phonetic detail and consequently how much memory traces there are for a given phonetic details. It is thus similar to the amount of exposure, although perceptual frequency refers in this case to the experience with vowel nasalization before the beginning of the experiment and not to the amount of exposure during training or test in the experiment. Assuming that high vowel nasalization corresponds to a low-frequency prime and that mid and low vowel nasalization corresponds to a high-frequency prime, there should be a greater increase of nasalization after exposure to high vowel nasalization than after exposure to non-high vowel nasalization in the exposure vowel. The increase in the exposure vowel might also be affected by ease of perception or ease of articulation. In these cases I either predicted the greatest increase after low vowel nasalization (ease of articulation) or after high vowel nasalization (ease of perception). The increase of nasalization in unexposed vowels differs depending on the theory: If there is a learning bias based on ease of perception, I will expect more generalizations to high vowels than to non-high vowels. If there is a learning bias based on ease of articulation, I will expect more generalizations to low vowels than to non-low vowels. If there is a learning bias for phonetically similar patterns, I will expect that low learners generalize more to mid than to high, that lower mid learners generalize more to low than to high and that high learners generalize more to mid than to low. For a summary of the predictions see Table 7. In this Table empty cells are those in which the exposure vowel would be the one most generalized to. In this case the effect of phonetics and the effect of exposure during the exposure phase of the experiment would not be separable. Cells including '/ ' are those in which the phonetic effect makes no prediction.

Table 7: Predictions for the productive task in the artificial language learning experiment testing vowel nasalization.

	Learning	**Generalizations of**		
		[a] -learners	**[ɛ] -learners**	**[i] -learners**
Perceptual frequency	[i] > [ɛ], [a]	/	/	/
Ease of perception	[i] > [ɛ], [a]	high > mid	high > low	
Ease of articulation	[a] > [ɛ], [i]		low > high	low > mid
Phonetic similarity	/	mid > high	low > high	mid > low

An increase of nasalization would be observable in a greater number of central vowels, in a more central vowel space and in lower A1−P0 values after exposure than before exposure.

As nasalization is not very common in German and as I do not have any representative baselines from former studies – except for the data collected for the production experiment in chapter 2.2 –, it may be the case, that an increase of nasalization might not be possible for German native speakers. Therefore, it might be the case that the participants compensate for this by decreasing the nasalization of those vowels that should profit from the smallest increase in nasalization, leaving the nasality levels of those that are expected to increase in nasalization as they are. A decrease of nasalization would be observable in a smaller number of central vowels, in a less central vowel space and in higher A1−P0 values after exposure than before exposure.

Thus, to test these predictions I checked how often the participants shifted their vowels to a more central vowel due to exposure, I plotted the mean vowel space for each group before and after exposure and I measured the difference in A1−P0 values before and after exposure.

3.2 Experiment

I will now describe how the experiment was designed by focusing on the stimuli, the procedure and the participants. After that I will present the analysis of the experimental results.

3.2.1 Method

Stimuli **Reading task.** For the reading task I created 31 German sentences which formed a little story. 15 of these sentences – every second sentence – contained a French loan word at the end of the sentence. These French loan words were: <balancieren> 'to balance', <revanchieren> 'to return a favor', <Parfum> 'perfume', <arrangiert> 'arranged', <Kartoffelgratin> 'potato gratin', <Fondant> 'fondant', <Bonbons> 'candy', <Interieur> 'interior', <Balkon> 'balcony', <Cousin> 'cousin', <Pensionär> 'retired person', <engagieren> 'to be involved in', <Branche> 'sector', <Waggon> 'wagon' and <Chance> 'chance'. In all of these items the vowel under investigation was followed by a nasal consonant, which causes the nasalization of the preceding vowel. These words were chosen as most of them are used in everyday communication. I therefore expected the participants to know these words – except for the words <Fondant>, which may only be known by persons who often bake, and <Interieur>, which is considered to be more formal than the other words.

In German these loan words can be pronounced in different ways – either with an oral vowel or with a nasal vowel. The choice of the oral vowel varies as well. In Table 8 there is an overview of the loan words used in the experiment with their pronunciation options for the vowels under investigation (printed in bold face) by German native speakers and their original French pronunciation. The German options can either be oral or they can be nasalized to a variable degree. Vowels marked with * (asterisk) were not observed in my data.

Table 8: French loan words with their German and their French pronunciation options.

French loan word	**German pronunciation**	**French pronunciation**
arr**a**ngiert	[a], [ɔ]	[ɑ̃]
bal**a**ncieren	[a], [ɔ]	[ɑ̃]
Br**a**nche	[a], [ɔ]	[ɑ̃]
Ch**a**nce	[a], [ɔ]	[ɑ̃]
engagieren	[a], [ɔ]	[ɑ̃]
Fond**a**nt	[a], [ɔ]	[ɑ̃]
rev**a**nchieren	[a], [ɔ]	[ɑ̃]
P**e**nsionär	[ɛ]	[ɑ̃]
Cous**i**n	[ɛ]	[ɛ̃]
Kartoffelgrat**i**n	[ɛ]	[ɛ̃]
Interieur	[i], [ɛ]	[ɛ̃]
Balk**o**n	[ɔ], *[oː]	[ɔ̃]
Wagg**o**n	[ɔ], *[oː]	[ɔ̃]
B**o**nbons	[ɔ]	[ɔ̃]
Parf**u**m	[y], [ɛ]	[œ̃]

Exposure phase. The stimuli for the perceptual exposure phase were the same as the stimuli in the perceptual training phase of the artificial language learning experiment on vowel nasalization (see chapter 2.4). Those stimuli were pseudowords from an artificial language, which consisted of singular, plural and diminutive forms. For details see chapter 2.4.1.

Procedure The experiment was divided into two reading tasks, a perceptual exposure phase and a forced-choice test. The experiment was scripted in Praat (Boersma & Weenink, 2019) and ran on a Windows laptop. It took place in an anechoic booth of the phonetics laboratory at Heinrich-Heine University Düsseldorf to avoid disturbing noise and lasted about 10 minutes. Participants listened to the auditory stimuli via headphones. For the recordings a microphone (phantom power 48V), a Sound Devices amplifier, a Marantz recorder and a Transtec computer were used. The sampling rate was 48 kHz.

There were three exposure groups: During exposure each group heard the nasalization of only one vowel – either [a], [ɛ] or [i]. For example, a member of the [a]-group never listened to any item with [ɛ] or [i]. The reading tasks were identical for all groups and included French loan words with high, mid and low nasal vowels in their German pronunciation.

At the beginning the participants were informed that there would be several phases and that there would be a short test about the content of a little story at the end of the experiment. The experiment started with the first reading task. Each sentence was displayed separately on the screen. After having read aloud one sentence the participants pressed a button and the next sentence appeared on the screen. While reading the participants' productions were recorded. After the first reading session the perceptual exposure to vowel nasalization followed. The perceptual exposure phase was identical to the training phase in the artificial language learning experiment testing vowel nasalization: The participants heard two repetitions of 48 auditory stimuli (16 singulars, 16 plural and 16 diminutives) in randomized order and saw images of fantasy animals on the screen. The subsequent reading task was the same as the one at the beginning of the experiment. The participants read aloud the same sentences again and their productions were recorded. The last task consisted of four multiple-choice questions about the story's content to disguise the experiment's aim. The responses to these questions were not analyzed.

Participants 52 adult native speakers of Northern German took part in the experiment. Due to technical problems or noise in the recordings, 45 recordings remained usable for the analysis. These recordings came from 33 female and 12 male participants (mean age: 22.8, range: 18-55). They were randomly assigned to one of the three experimental groups. There were 15 participants in each group. No one reported knowledge of a language that uses nasalized vowels distinctively. All participants had normal or corrected-to-normal vision, no hearing problems and did not suffer from hoarseness during recording. None of them had participated in the previous experiments (production experiment, perception experiment, artificial language learning experiment). Participants were given a small expense allowance for their participation. None was an advanced student of linguistics but all of them were students or researchers at Heinrich-Heine University Düsseldorf. This was done to control for the educational background which can influence the degree of nasalization according to Greisbach (2003) and Laeufer (2010). Prior to the experiment each participant filled out a consent form and a questionnaire with which data about the participants' age, sex, native language, foreign language skills, the region they had been grown up and education was collected.

3.2.2 Results

Prior to analysis the data had to be prepared, which means that textgrids were created in Praat (Boersma & Weenink, 2019) in which the loan words as well as the vowel under investigation were labeled. The selection and the labeling of the vowel had to be consistent across all recordings so that the measurements of HNR, of F1 and F2 and of A1−P0 were not confounded by inconsistent vowel boundaries. The vowels' starting point and the vowels' end point was always a zero-crossing point. The vowel boundaries were set based on visual inspections of the wave form and the spectrogram. The first quasiperiodic wave of the vowel was chosen as starting point and the last quasiperiodic wave was chosen as end point of the vowel. This selection was in agreement with the visible formants in the spectrogram. This procedure was exactly the same as in the production experiment described in chapter 2.2 and is illustrated in Figure 8.

The original data set consisted of 1350 vowels (15 words × 45 participants × 2 repetitions). 14 words had to be excluded due to mispronunciations and 3 words had to be excluded due to incomplete recordings. As nasality has similar acoustic features as creaky voice (Zhang, 2015), I also had to exclude all words (n = 195) produced with creaky voice. Similar to the procedure in the production experiment in chapter 2.2 I decided to measure HNR with Praat (Boersma & Weenink, 2019) to determine which of the recordings sound creaky. As sounds with lower HNR are creakier, I ordered the stimuli according to their HNR-value. A phonetically trained linguist and I listened independently from each other to the ones with the lowest HNR. In addition to that we checked the distance between the pulses visually because irregularly spaced pulses are a further hint to creaky voice. By doing so, both listeners labeled all sounds with an HNR lower than 7 dB as creaky, which was then chosen as cut-off point. In Figures 15 and 16 you can see the difference between recordings with creak (see Figure 15) and without creak (see Figure 16). Figure 15 shows the oscillogram (above) and spectrogram (below) of [ɔ] in <Bonbons> with an HNR of 3.42 dB (speaker-04: male, recording 2). Irregular spaced pulses (vertical blue lines) – damped in comparison to those in Figure 16 – and an irregular and low F0 (horizontal blue line, mean 114.2 Hz) are further observable acoustic indicators for the presence of creaky voice in this recording. Figure 16 shows the oscillogram (above) and spectrogram (below) of [ɔ] in <Bonbons> with an HNR of 16.29 dB (speaker-50: male, recording 2). Regular spaced pulses (vertical blue lines) and a regular and higher F0 (horizontal blue line, mean 145.5 Hz) than in Figure 15 are further acoustic indicators for the absence of creaky voice in this recording. The remaining words (n = 1138) were further analyzed.

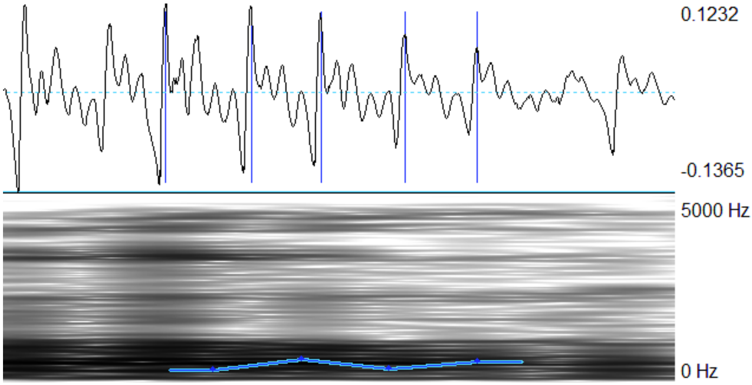

Figure 15: Example of a recording with creak.

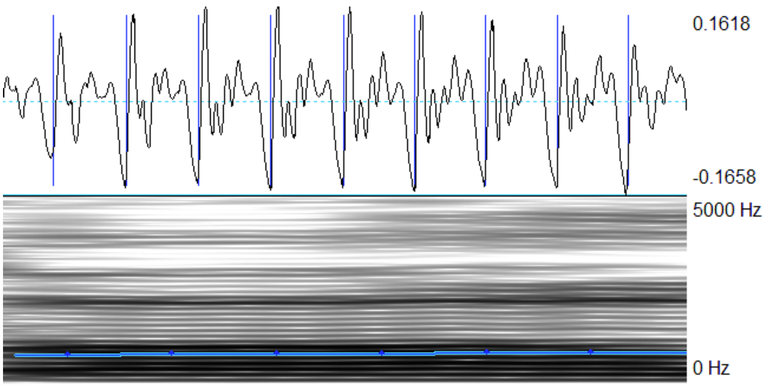

Figure 16: Example of a recording without creak.

First of all, I decided which vowel was pronounced in each word with multiple pronunciation options. This was done because several loan words can be pronounced with a more peripheral vowel (low [a] or high [i] and [y]) – meaning a better adaption to German – or with a more central vowel (mid [ɔ] and [ɛ]) – meaning a better imitation of the French nasalization. The reason for this might be that due to the nasalization the (perceptual) vowel height is lowered for high vowels and raised for low vowels – leading to a smaller and more centralized vowel space in nasal vowels than in oral vowels (Beddor, 1993; Beddor et al., 1986; Fant, 1960; Hecker, 1962; Wright, 1975) as described in chapter 3.1. The word <balancieren> can, for example, either be pronounced as [bal**a**ŋsiːʁən] (with low [a] before the nasal consonant) or as [bal**ɔ**ŋsiːʁən] (with mid [ɔ] before the nasal consonant), whereas the second option is considered to be a better imitation of the French nasalization. As nasalization is relatively uncommon in German, a shift in vowel quality may be a hint at a more nasal pronunciation by German native speakers being not familiar with French or any language with phonemic nasalization. Second, I plotted the pronounced vowels of all loan words for each recording time separately in the vowel space. This was done to find smaller shifts in vowel quality not detected by listening. Third, I measured nasality acoustically by means of the A1−P0 value to objectively compare the amount of nasality in the items depending on exposure group and vowel height.

Centralization of vowels As different pronunciations of French loan words are possible in German, I first of all made an overview of the vowels pronounced in each loan word in the first recording and in the second recording. In Table 9 you can see for each group the percentages of pronunciation options in each item with variable pronunciation before and after exposure. The second options [ɔ] and [ɛ] are considered to be better imitations of vowel nasalization than the first options [a], [y] and [i]. A positive value in *mean centralization* means an increase of the central option after exposure. The overall change refers to the mean centralization across all exposure groups for each word.

Several persons were involved in the preparation of the data. To decide which of the vowel options was articulated a student assistant and I listened independently of each other to the recordings and labeled the vowels. When there was any discrepancy, a third phonetically trained person was asked to decide which vowel was articulated. After some months I listened to the recordings again and checked the labeling once more. Again, when there was any discrepancy, a phonetically trained person was asked for advice. I decided to rely mainly on the auditory impression when labeling the vowel options as the measurement of formant frequencies is highly speaker and gender dependent so that I would have needed more data per speaker. Moreover, as formant frequencies shift due to the nasalization, they would not have been reliable either.

Table 9: Percentages of pronunciation options in each item with variable pronunciation before and after exposure.

French loan word	**Exposure group**	**Vowel before exposure**	**Vowel after exposure**	**Mean centralization**
arrangiert	[a]	57% [a], 43% [ɔ]	27% [a], 73% [ɔ]	+30%
	[ɛ]	64% [a], 36% [ɔ]	56% [a], 44% [ɔ]	+8%
	[i]	29% [a], 71% [ɔ]	33% [a], 77% [ɔ]	+6%
				overall: +11%
balancieren	[a]	64% [a], 36% [ɔ]	50% [a], 50% [ɔ]	+14%
	[ɛ]	57% [a], 43% [ɔ]	57% [a], 43% [ɔ]	0%
	[i]	71% [a], 29% [ɔ]	69% [a], 31% [ɔ]	+2%
				overall: +5%
Fondant	[a]	27% [a], 73% [ɔ]	36% [a], 64% [ɔ]	−9%
	[ɛ]	25% [a], 75% [ɔ]	15% [a], 85% [ɔ]	+10%
	[i]	38% [a], 62% [ɔ]	31% [a], 69% [ɔ]	+7%
				overall: +4%

Continuation of the Table on the next page.

Table 9: Continuation of Table 9.

Branche	[a]	7% [a], 93% [ɔ]	7 % [a], 93% [ɔ]	0%
	[ɛ]	0% [a], 100% [ɔ]	0% [a], 100% [ɔ]	0%
	[i]	0% [a], 100% [ɔ]	0% [a], 100% [ɔ]	0%
				overall: 0%
Chance	[a]	7% [a], 93% [ɔ]	0% [a], 100% [ɔ]	+7%
	[ɛ]	8% [a], 92% [ɔ]	8% [a], 92% [ɔ]	0%
	[i]	7% [a], 93% [ɔ]	0% [a], 100% [ɔ]	+7%
				overall: +5%
revanchieren	[a]	7% [a], 93% [ɔ]	7% [a], 93% [ɔ]	0%
	[ɛ]	7% [a], 93% [ɔ]	7% [a], 93% [ɔ]	0%
	[i]	13% [a], 87% [ɔ]	13% [a], 87% [ɔ]	0%
				overall: 0%
engagieren	[a]	83% [a], 17% [ɔ]	50% [a], 50%[ɔ]	+33%
	[ɛ]	86% [a], 14% [ɔ]	100% [a], 0% [ɔ]	−14%
	[i]	89% [a], 11% [ɔ]	86% [a], 14% [ɔ]	+3%
				overall: +2%
Parfum	[a]	67% [y], 33% [ɛ]	67% [y], 38% [ɛ]	+5%
	[ɛ]	92% [y], 8% [ɛ]	83% [y], 17% [ɛ]	+9%
	[i]	69% [y], 31% [ɛ]	50% [y], 50% [ɛ]	+19%
				overall: +10%
Interieur	[a]	100% [i], 0% [ɛ]	100% [i], 0% [ɛ]	0%
	[ɛ]	100% [i], 0% [ɛ]	78% [i], 22% [ɛ]	+22%
	[i]	89% [i], 11% [ɛ]	100% [i], 0% [ɛ]	−11%
				overall: +5%

Overall, the pronunciation was highly variable, depending on both, word and speaker. Comparing the options chosen in recording 1 in Table 9, you can see that there was a very strong preference for one vowel over the other vowel in some words (e.g. <Branche>, <Chance>, <Interieur>), whereas this preference was not that strong in other words (e.g. <arrang-

iert>, <balancieren>). Besides, the options chosen in recording 1 varied for the same words (e.g. <arrangiert>, <Parfum>) between groups. As this recording was made prior to exposure and as participants were randomly assigned to one of the groups, this variation could not be attributed to the experimental set-up. The variation between groups in recording 2 – after exposure to nasalization of different vowel heights –, however, could be attributed to the experimental set-up. I will explore this variation in detail below. For this analysis chances below 5% are considered to be random and called *no changes*.

In general, a change towards the peripheral pronunciation was rare. It could be observed in <Interieur> after [i]-exposure, in <engagieren> after [ɛ]-exposure and in <Fondant> after [a]-exposure. In two of the three cases, the peripheral option reached 100% (<Interieur> and <engagieren>). No change at all could be observed in <Branche> and <revanchieren> for all groups, in <balancieren> after [ɛ]- and [i]-exposure, in <Chance> after [ɛ]-exposure, in <Interieur> and <Parfum> after [a]-exposure and in <engagieren> after [i]-exposure. In some of these words the central option was already the preferred one in recording 1 and yielded 100% (e.g. <Branche> in the [ɛ]- and [i]-group). Thus, an increase in the central options was not always possible. Most of these *no change* words preferred the central options (n = 9), whereas three of them preferred the peripheral options. A change towards the central pronunciation took place most often and it could be observed in <arrangiert> for all groups, in <balancieren> after [a]-exposure, in <Fondant> after [ɛ]- and [i]-exposure, in <Chance> after [a]- and [i]-exposure, in <Parfum> after [ɛ]-and [i]-exposure, in <Interieur> after [ɛ]-exposure and in <engagieren> after [a]-exposure.

The exposure groups patterned together in different ways. In <Branche> and <revanchieren> all groups showed the same tendency: The central option was more common in both recordings for all groups and there was no change after exposure. In <arrangiert> all groups behaved similar as well: They all increased the central option. All other words, however, displayed different behaviors across groups, with either all groups showing a different tendency, or the [i]-group and the [ɛ]-group patterning together or the [a]-group and the [i]-group patterning together. The [a]-group and the [ɛ]-group never patterned together. In <Interieur> and <engagieren> all groups showed a different tendency. <Interieur> shifted to the central option after [ɛ]-exposure, to the peripheral option after [i]-exposure and not at all after [a]-exposure, whereas <engagieren> shifted to the central option after [a]-exposure, to the peripheral option after [ɛ]-exposure and not at all after [i]-exposure. [i]-Group and [ɛ]-group patterned together in <balancieren>, <Fondant> and <Parfum>. For <balancieren> there was a shift towards the central option after [a]-exposure, but no change at all after [i]- and [ɛ]-exposure. For <Fondant> and <Parfum> one could observe a shift to the central option after [ɛ]- and [i]-exposure, but a change to the peripheral option (<Fondant>) or no change at all (<Parfum>) after [a]-exposure. In <Chance> the [a]-group and the [i]-group patterned together. This word was shifted to the central option after [a]- and [i]-exposure but its pronunciation was not changed at all after [ɛ]-exposure.

Overall, the [a]-group shifted their pronunciations more often towards the central option than the [ɛ]- and the [i]-group (see Table 10).

Table 10: Mean shift to the central vowel pronunciation for each exposure group.

Exposure group	**Overall**	**[a] - [ɔ]**	**[y] - [ɛ]**	**[i] - [ɛ]**
[a]	+11%	+10%	+5%	0%
[ɛ]	+3%	+1%	+9%	+22%
[i]	+3%	+1%	+19%	−11%

Moreover, it could be observed that after [a]-exposure the shift was stronger for the low vowel than for the high vowels, whereas after [ɛ]- and [i]-exposure the shift was stronger for high vowels (although this shift was not always towards the central option). Note that the shifts from [y] to [ɛ] and from [i] to [ɛ] were each based on one word only (<Parfum> and <Interieur>) so that these statements may be inconclusive. The same is true for the further similar analyses.

To test whether the observed shifts in vowel quality reached significance, I calculated a linear regression model with PRODUCED VOWEL (central and peripheral) as dependent variable and RECORDING (before exposure and after exposure) as independent variable in R (R Core Team, 2015) (R syntax: lm(PRODUCED VOWEL ~ RECORDING)) for each of the three exposure groups (low, mid and high) separately. As PRODUCED VOWEL and RECORDING are both nominal variables, I dummy coded them (central vowel = 0 and peripheral vowel = 1, before exposure = 0 and after exposure = 1). The results show that neither of the observed changes reached significance: The choice of the vowel did not depend on whether the participants where exposed to nasalization before or not.

There was no significant difference in the choice of vowel depending on exposure for the exposure vowel [a] (Est. = 0.1052, SE = 0.0647, t = 1.627, p = 0.105), for the exposure vowel [ɛ] (Est. = 0.0337, SE = 0.0670, t = 0.504, p = 0.615) or for the exposure vowel [i] (Est. = 0.0246, SE = 0.0648, t = 0.380, p = 0.704) in general.

Then I checked, whether there were significant differences for the different pronunciation options of the French loan words. To do this, I created three further subsets for each exposure group depending on the pronunciation options of the loan words ([a]-[ɔ], [i]-[ɛ], [y]-[ɛ]) and ran the same model described above on these three subsets. There was no significant difference in the choice of vowel for the pair [a]-[ɔ] depending on exposure to the vowel [a] (Est. = 0.1028, SE = 0.0662, t = 1.553, p = 0.122), to the vowel [ɛ] (Est. = 0.0032, SE = 0.0704, t = 0.045, p = 0.964) or to the vowel [i] (Est. = 0.0148, SE = 0.0683, t = 0.217, p = 0.829). There was no significant difference in the choice of vowel for the pair [i]-[ɛ] depending on exposure to the vowel [a] (Est. = 0.0833, SE = 0.2228, t = 0.374, p = 0.712), to the vowel [ɛ] (Est. = 0.2636, SE = 0.1861, t = 1.417, p = 0.173) or to the vowel [i] (Est. = 0.0154, SE = 0.1920, t = 0.080, p = 0.937). There was no significant difference in the choice of vowel for the pair [y]-[ɛ] depending on exposure to the vowel [a] (Est. = 0.0513, SE = 0.1882, t = 0.273, p = 0.787), to the vowel [ɛ] (Est. = 0.2024, SE = 0.1560, t = 1.297, p = 0.207) or to the vowel [i] (Est. = 0.1429, SE = 0.2002, t = 0.714, p = 0.482).

I also tested whether there were significant differences in the choice of vowel for each word separately depending on the height of the exposure vowel. To do this, I created further subsets for each exposure group depending on the loan word itself and ran the same model described above on these subsets. Neither of the models showed a significant difference in vowel quality depending on exposure. The reasons for this might either be the great variability or the small number of observations. The latter has also been the reason why I was not able to perform a chi-squared-test, which is most suitable for nominal data.

As the French loan words used in the experiment occur with different frequency in German, I checked whether this can explain the patterning. I checked the type frequency of the loan words in the German dlex corpus (Heister, Würzner, Bubenzer, Pohl, Hanneforth, Geyken & Kliegl, 2011). This corpus is the reference corpus of the German language of the 20th century and it is available online: http://www.dlexdb.de/query/kern/typposlem/. It contains four almost equally distributed genres (28% fiction, 27% newspapers, 23% scientific publications and 21% functional literature) with 2.3 millions types. The type frequency[8] of each French loan word used in the reading task in the German language based on the German dlex corpus (Heister et al., 2011) and the mean shift to the central option in the

[8]The type frequency of 1 for *Fondant* means that there is one occurrence of *Fondant* in the 2.3 millions types in the corpus.

recordings (compare Table 9) are displayed in Table 11.

Table 11: Type frequency of the French loan words in German and mean shift to the central option in the recordings.

French loan word	**Type frequency**	**Mean overall centralization**
Fondant	1	+4%
(Kartoffel)Gratin	5	
Pensionär	41	
Interieur	84	+5%
revanchieren	104	0%
balancieren	124	+5%
Cousin	159	
arrangiert	210	+11%
engagieren	241	+2%
Bonbons	244	
Waggon	430	
Parfum	462	+10%
Branche	598	0%
Balkon	905	
Chance	2490	+5%

What can be seen in Table 11 is that the as *no changes* considered words <Fondant>, <Interieur>, <revanchieren> and <balancieren> as well as the *no change* words <Branche> and <Chance> are either the ones with the lowest type frequencies (1-124) or the ones with the highest type frequencies (598-2490). Those loan words which were most affected by the exposure to nasalization (<arrangiert> and <Parfum>) are in between with type frequencies between 210 and 462. Thus, the pronunciation of loan words with an intermediate type frequency can be changed due to auditory exposure, whereas that of loan words with a low or high type frequency cannot. The only exception here is <engagieren>, which is a *no change* word as well but it is in between <arrangiert> and <Parfum>. However, in this case one has to keep in mind that <engagieren> is highly variable across groups: Participants being exposed to [a] tended to increase the nasalization (+33%), whereas those being exposed to [ɛ] tended to decrease the nasalization (−14%). Hence, this word was indeed affected by the exposure to nasalization.

The type frequency data might also explain the irregularities, e.g. mispronunciations and pauses, observed in the recordings of the words <Interieur> and <Fondant>. Some participants were not sure how to pronounce them due to unfamiliarity with these words. This unfamiliarity can either be due to the uncommon orthography for a German word with <ieur> at the end of the word or due to the low type frequency of a more specialized word (cake decoration). As a consequence, the participants stuck to the peripheral option – the better adaption to German – almost to 100%. <Chance> and <Branche> are highly frequent and are those words which were produced most often with the central option – the better imitation of French nasalization – in recording 1 and in recording 2. Hence, these loan words seem to be integrated into the German vocabulary with its original pronunciation.

The shift in vowel quality can also be very subtle, making it impossible to hear. Moreover, it might be the case that the vowels also shifted in the loan words without multiple vowel options. Therefore I plotted the vowels in the acoustic vowel space based on F1 and F2 with the help of R (R Core Team, 2015) and the corresponding R packages ggplot2 (Wickham, 2009) and dplyr (Wickham, François, Henry & Müller, 2019). With Praat (Boersma & Weenink, 2019) and the Nasality Automeasure script (Styler & Scarborough, 2017) I measured F1- and F2-values at three different time points for each vowel in a French loan word relating to 50%, 75% and to the end of the vowel interval. This was done as the vowels under investigation always occur before a nasal consonant so that I expected the greatest amount of nasality in the vowel portion next to the nasal consonant. The aim of the plotted vowel spaces was to see whether there were shifts from more peripheral vowels before exposure to more central vowels after exposure. For these plots I used the 3414 data from all loan words (1138 words × 3 time points). The plots are shown separately for each exposure group in Figures 17, 18 and 19. In these Figures the vowel space before and after exposure to vowel nasalization is displayed. Gray vowel labels marked with *1* form the vowel space before exposure and are labeled with vowel_1, e.g. i_1. Red vowel labels marked with *2* form the vowel space after exposure and are labeled with vowel_2, e.g. i_2. The ellipses around the vowels correspond to ±1 SD and are drawn with different types of lines, e.g. solid lines for the vowel [i] or dotted lines for the vowel [ɛ].

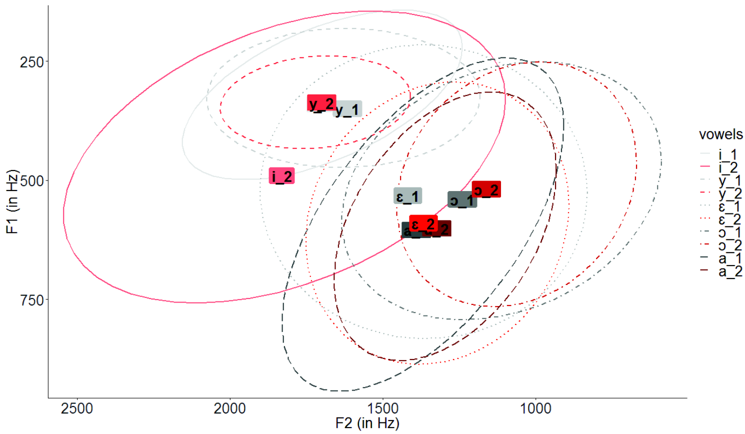

Figure 17: Vowel space before (gray, 1) and after (red, 2) exposure to low vowel nasalization. Ellipses correspond to ±1 SD.

The plot in Figure 17 shows that [a]-exposure did not lead to a centralized vowel space. None of the vowels was shifted towards a more central space after exposure. When looking at the three vowel shift-pairs, one can observe a shift from [a] towards [ɔ] due to exposure, which was achieved by a minimal lowering of F2. However, there was no lowering of F1. [i] was shifted in the direction of [ɛ] due to exposure by raising F1. However, there was no lowering of F2. [y], however, was not shifted towards [ɛ] after [a]-exposure.

In Figure 18 one can observe a centralization effect for all vowels, except [a], which led to a centralized vowel space after [ɛ]-exposure. When looking at the three vowel pair-shifts, one can observe that [i] was shifted towards [ɛ] due to exposure by lowering of F2 and by raising of F1. [y] was minimally shifted towards [ɛ] due to exposure by lowering F2. However, there was no raising of F1. After [ɛ]-exposure there was no shift from [a] towards [ɔ].

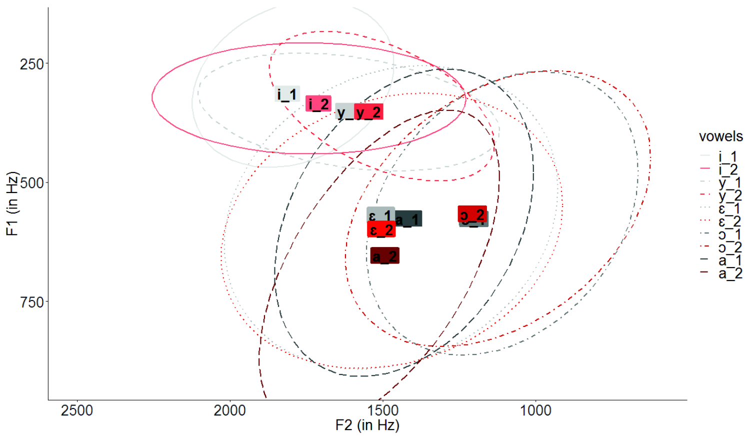

Figure 18: Vowel space before (gray, 1) and after (red, 2) exposure to mid vowel nasalization. Ellipses correspond to ±1 SD.

The vowel space in Figure 19 shows that all vowels, except [a] and [ɔ] were somewhat more central after [i]-exposure. Looking at the three vowel pair-shifts, one can observe that [y] was shifted towards [ɛ] due to exposure by a raising of F1. However, there was no lowering of F2. [i] was shifted towards [ɛ] due to exposure by lowering F2. However, F1 was minimally lowered – not raised – as well, so that there was no shift towards [ɛ] based on F1 after [i]-exposure. A similar picture emerges for the vowel [a]: There was a shift towards [ɔ] due to a minimal lowering of F2. There was, however, no lowering of F1 in this case.

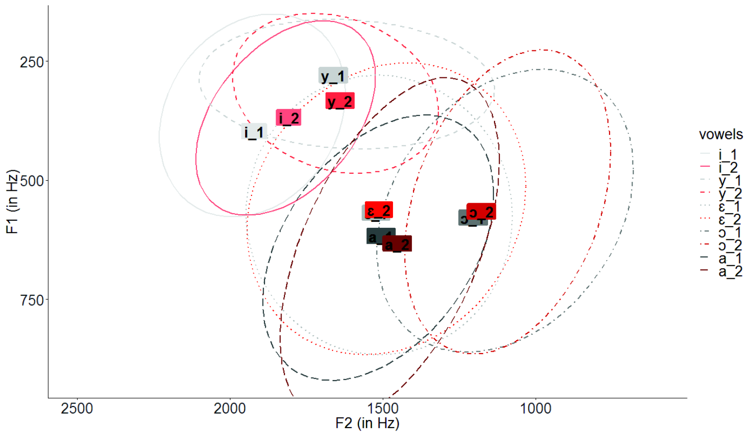

Figure 19: Vowel space before (gray, 1) and after (red, 2) exposure to high vowel nasalization. Ellipses correspond to ±1 SD.

Comparing the vowel spaces of the three exposure groups with each other, one can observe that the shift from low [a] to more central [ɔ] was strongest after exposure to low vowel nasalization. The shift was not present after exposure to mid vowel nasalization and only partially fulfilled after exposure to high vowel nasalization. Moreover, the shift from high [y] to more central [ɛ] was strongest after exposure to high and mid – non-low – vowel nasalization and not present at all after exposure to low vowel nasalization. The shift from high [i] to more central [ɛ], however, was only present after exposure to mid vowel nasalization. After exposure to low and high vowel nasalization there was only a shift of one formant in the direction of more central [ɛ], whereas the other formant was shifted further away from [ɛ]. These results, which are based on visual inspection, are in line with the results based on listening to the recordings presented in Table 10. Note that there was a large variability among the F1- and F2-values (see size of the ellipses in the vowel spaces) so that all shifts described here remain tendencies. The same is true for the results in Table 10 which did not reach statistical significance.

I also calculated the vowel space area (VSA) in R (R Core Team, 2015) with the help of the R package PhonR (McCloy, 2016) to see whether the exposure to nasalization caused a centralization of the vowel space as proposed by e.g. House (1957) and the visual inspection. Specifically, I calculated the area of the convex hull (F2 $\times$ F1), whose calculation is based on all vowel tokens, for each exposure group (low, mid and high) separately (R syntax: `convexHullArea(F1, F2, group = recording)`). The vowel space area was smaller, which means a more centralized vowel space, after exposure to high vowel nasalization than before exposure to high vowel nasalization (before: 2630725, after: 2621236, difference: -9489, -0.4%). The same was true for exposure to mid vowel nasalization (before: 2304583, after: 1758632, difference: -545951, -24%). However, the vowel space area was bigger after exposure to low vowel nasalization than before exposure (before: 2225629, after: 2270819, difference: $+45190$, $+2\%$). This suggests a greater centralization effect after exposure to [ɛ] than after exposure to [i] and a de-centralization effect after exposure to [a]. This calculation confirms the observed patterns in the plotted vowel spaces in Figures 17-19.

Acoustic nasalization: A1−P0 With Praat (Boersma & Weenink, 2019) and the Nasality Automeasure script (Styler & Scarborough, 2017) I measured A1−P0 values at three different time points for each vowel in the French loan word relating to 50%, 75% and to the end of the vowel interval. This was done as the vowels under investigation always occur before a nasal consonant so that I expected the greatest amount of nasality in the vowel portion next to the nasal consonant. Due to problems with the measurement indicated by the script (A1 = P0 (n = 478), Shallow (n = 140), LoPitch (n = 42), HighF1 (n = 24), HarmDev (n = 14), Crash-PulseBegin (n = 4), HiPitch (n = 2)), only 2710 data points of the original 3414 data points (1138 vowels × 3 time points) were analyzable. A description of most of these errors can be found in chapter 2.2.2. Crash-PulseBegin is due to an unconfirmed pulse length (for details see Styler & Scarborough (2017)).

Overall one could observe only a small increase of the amount of nasality. The mean A1−P0 value at recording 1 was −2.41 dB and it decreased by −0.14 dB due to exposure to nasalization so that it was −2.54 dB in recording 2. These values show that German native speakers nasalize vowels in French loan words more than vowels in nasal contexts of pseudowords, which conform to the German phonotactics as a comparison with the A1−P0 values from the production experiment (see chapter 2.2) reveals. In the pseudowords high vowels reached an A1−P0 value of 1.10 dB, mid vowels of 0.33 dB and low vowels of 0.86 dB. These higher values indicate a less amount of nasality.

Before exposure the amount of nasality was greatest in low vowels in all groups ([a]-group: −4.80 dB, [ɛ]-group: −5.48 dB, [i]-group: −3.33 dB), followed by mid ([a]-group: −2.52 dB, [ɛ]-group: −2.27 dB, [i]-group: −1.46 dB) and high vowels ([a]-group: −3.25 dB, [ɛ]-group: 0.14 dB, [i]-group: −1.46 dB) (see Figures 20-22). This difference, however, was not significant as three separate linear mixed-effects model – one for each group before exposure – calculated in R (R Core Team, 2015) with the R packages lme4 (Bates et al., 2015) and lmerTest

(Kuznetsova et al., 2018) show. In the models which fitted the data best (as assessed by backward stepwise elimination (Baayen, 2008)) A1−P0 was the independent variable and VOWEL HEIGHT (high, mid and low) was the dependent variable. PARTICIPANT and ITEM served as random intercepts (R syntax: lmer(A1−P0 ~ VOWEL HEIGHT + (1|PARTICIPANT) + (1|ITEM))). There was no significant difference in the amount of nasality before exposure to low vowel nasalization between high and mid vowels (Est. = −1.322, SE = 1.550, df = 139.081, t = −0.853, p = 0.395), between high and low vowels (Est. = −1.067, SE = 1.8190, df = 150.240, t = −0.587, p = 0.558) and between mid and low vowels (Est. = 0.2548, SE = 0.9891, df = 313.951, t = 0.258, p = 0.797). There was also no significant difference in the amount of nasality before exposure to mid vowel nasalization between high and mid vowels (Est. = −2.2574, SE = 2.2878, df = 14.425, t = −0.987, p = 0.340), between high and low vowels (Est. = −3.1500, SE = 2.4074, df = 17.283, t = −1.308, p = 0.208) and between mid and low vowels (Est. = −0.8926, SE = 0.9568, df = 325.783, t = −0.933, p = 0.352). The same is true for the exposure to high vowel nasalization. Before exposure to high vowel nasalization there was no significant difference in the amount of nasality between high and mid vowels (Est. = −0.7941, SE = 1.5101, df = 81.404, t = −0.526, p = 0.600), between high and low vowels (Est. = −1.2205, SE = 1.6883, df = 86.346, t = −0.723, p = 0.472) and between mid and low vowels (Est. = −0.4264, SE = 0.9240, df = 216.219, t = −0.462, p = 0.645).

I was also interested in whether there were differences in the amount of nasality in vowels of different height in the three groups after exposure depending on the nasal vowel height heard during exposure. I therefore calculated a separate linear mixed-effects model in R (R Core Team, 2015) with the R packages lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2018) for each vowel height in each group. In the models which fitted the data best (as assessed by backward stepwise elimination (Baayen, 2008)) the A1−P0 value was the independent variable and RECORDING (before exposure and after exposure) was the dependent variable. PARTICIPANT and ITEM served as random intercepts (R syntax: lmer(A1−P0 ∼ RECORDING + (1|PARTICIPANT) + (1|ITEM))).

With the R packages ggplot2 (Wickham, 2009) and Rmisc (Hope, 2013) I plotted the mean A1−P0 values depending on vowel height (high, mid and low) and recording (before exposure and after exposure) for each exposure group separately (see Figures 20, 21 and 22). Higher bars (= lower values) correspond to greater acoustic nasality.

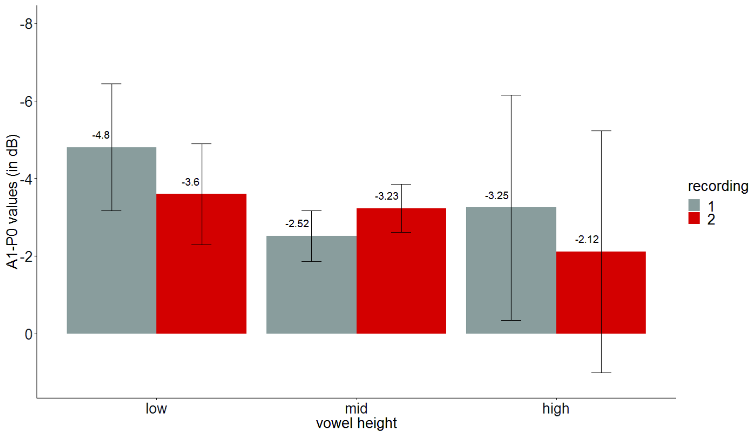

Figure 20: A1−P0 values (in dB) ± 1.96 SE before (gray) and after (red) exposure to nasalization of [a].

The three linear mixed-effects models – one for each vowel height – for the [a]-group show, that there was no significant difference in the amount of nasality between recording 1 (low: −4.80 dB, mid: −2.52 dB) and recording 2 (low: −3.60 dB, mid: −3.23 dB) for low (Est. = 1.2370, SE = 1.0130, df = 113.861, t = 1.222, p = 0.224) and mid vowels (Est. = −0.3040, SE = 0.3548, df = 700.575, t = −0.857, p = 0.392). There was a marginally significant difference in the amount of nasality between recording 1 (−3.25 dB) and recording 2 (−2.12 dB) for high vowels (Est. = 2.1420, SE = 1.2290, df = 44.424, t = 1.744, p = 0.088). An overview of the results for the [a]-group can be found in Figure 20.

The three linear mixed-effects models – one for each vowel height – for the [ɛ]-group show, that there was no significant difference in the amount of nasality between recording 1 (low: −5.48 dB, mid: −2.27 dB, high: 0.14 dB) and recording 2 (low: −6.12 dB, mid: −3.00 dB, high: −0.43 dB) for low (Est. = 0.1044, SE = 0.7152, df = 123.903, t = 0.146, p = 0.884), mid (Est. = −0.6209, SE = 0.3861, df = 663.259, t = −1.608, p = 0.108) and high vowels (Est. = −1.2384, SE = 0.9647, df = 58.686, t = −1.284, p = 0.204). An overview of the results for the [ɛ]-group can be found in Figure 21.

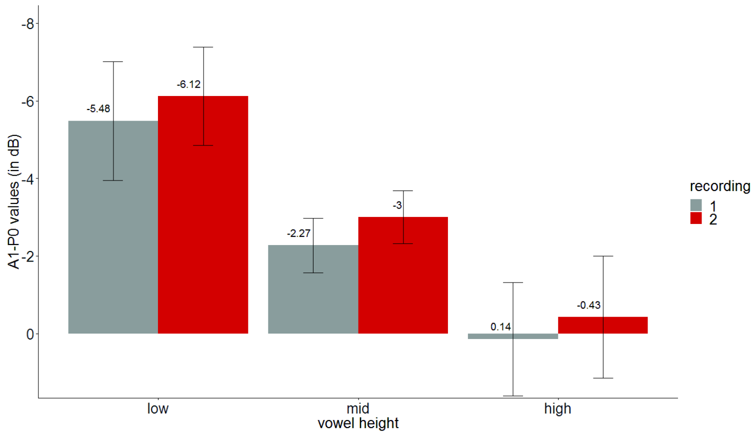

Figure 21: A1−P0 values (in dB) ± 1.96 SE before (gray) and after (red) exposure to nasalization of [ɛ].

The three linear mixed-effects models – one for each vowel height – for the [i]-group show, that there was a significant decrease of nasality between recording 1 (low: −3.33 dB, mid: −1.46 dB, high: −1.46 dB) and recording 2 (low: −1.82 dB, mid: −0.74 dB, high: −3.59 dB) in low vowels (Est. = 1.7024, SE = 0.7461, df = 135.200, t = 2.282, $p < 0.05$). There was, however, no significant difference in the amount of nasality between recording 1 and recording 2 in mid (Est. = 0.6839, SE = 0.4094, df = 678.800, t = 1.670, p = 0.095) and high vowels (Est. = −0.1471, SE = 1.3258, df = 55.000, t = −0.111, p = 0.912) for the [i]-group. An overview of the results for the [i]-group can be found in Figure 22.

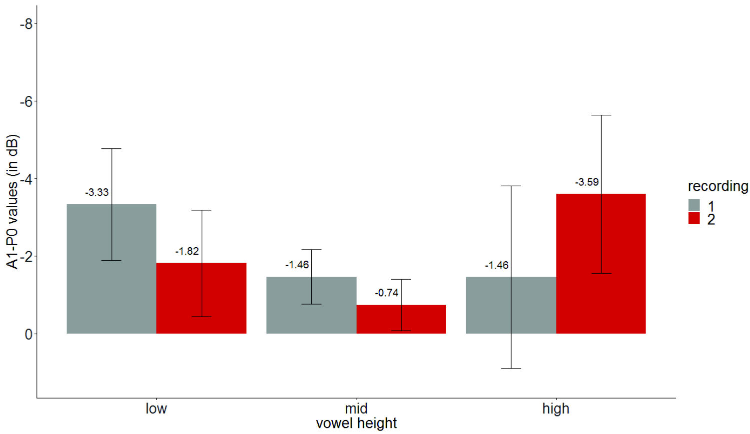

Figure 22: A1−P0 values (in dB) ± 1.96 SE before (gray) and after (red) exposure to nasalization of [i].

3.3 Discussion

The experiment shows that native speakers of German adjust their degree of nasalization after auditory exposure to a vowel nasalization pattern. This adjustment, however, is quite variable and depends on the vowel height of the exposure vowel.

The pronunciation of nasal vowels in French loan words by German native speakers is in general highly variable and seems to depend on the word itself, its lexical frequency and the speaker. It seems as if some words have a preferred pronunciation by several speakers which either can (e.g. <Parfum>, <arrangiert>) or cannot be changed (e.g. <revanchieren>) by auditory exposure to a specific vowel height. When having listened to nasal vowels – independent of the vowel's height – the pronunciation was most often shifted towards the more central option, which is an appropriate imitation of French nasalization. A shift towards the more peripheral option was rare.

Depending on vowel height of the exposure vowel a shift towards the more central option occurred more often when the exposure vowel was low than when it was mid or high. However, these observed shifts did not reach significance. There was, however, the tendency that exposure to nasalization centralized the vowel (better imitation of French vowel nasalization) and that this effect was stronger for those speakers who had listened to nasal vowels which were easy to articulate (low vowels). Moreover, it could be observed that the shift from low [a] to more central [ɔ] was strongest after exposure to low vowel nasalization. This shift was not present after exposure to mid vowel nasalization and only partially fulfilled after exposure to high vowel nasalization. The shift from high [y] to more central [ɛ] was strongest after exposure to high and mid – non-low – vowel nasalization and not present at all after exposure to low vowel nasalization. These observed shifts were minimally and were just displayed for data visualization. Nevertheless, based on these observations, it seems as if the vowel height of the exposure vowel plays a role in these shifts, with exposure to low vowel nasalization shifting low vowels to mid and exposure to non-low vowel nasalization shifting non-low vowels to mid. However, for the shift from high [i] to more central [ɛ] this is not that clear: Only after exposure to mid vowel nasalization there was a shift of both formants in the predicted direction. After exposure to low and high vowel nasalization there was only a shift of one formant in the direction of more central [ɛ], whereas the other formant was shifted in the opposite direction – further away from [ɛ].

Independent of these vowel specific shifts, the centralization effect was only present after exposure to non-low vowel nasalization. This suggests that the production of nasal vowels is not that heavily affected by auditory exposure to low vowel nasalization than by auditory exposure to high or mid vowel nasalization. Moreover, as the centralization effect was not present for [a] (and [ɔ]) it seems that the production of non-low vowels and back vowels is not that much affected by auditory exposure than the production of high, mid and front vowels. These results are in line with the calculations of the vowel space area, which show a small de-centralization effect after exposure to low vowel nasalization and a centralization effect after exposure to mid vowel nasalization. After exposure to high vowel nasalization there seems to be neither a centralization nor a decentralization effect.

Carignan (2018) analyzed six different languages and found the following patterns by measurements of nasalance (ratio of acoustic energy coming from the nose among acoustic energy coming from the nose and the mouth (Fletcher, Sooudi & Frost, 1974)) and ultrasounds, which are only partly in agreement with the centralization effect: (1) raising of F1 and F2 of [i], (2) lowering of F1 of non-high vowels and (3) lowering of F2 of non-front vowels in nasal context compared to oral context. From these statements only the second one fully agrees with the centralization effect. The first statement agrees with the centralization effect in the raising of F1 but not in the raising of F2, and the third statement does not fully agree with the centralization effect, as well. My data confirms the first statement for [a]-learners, the second statement for [ɔ] in all groups (only minimally) and for [ɛ] in [i]-learners, and the third statement for [a]- and [i]-learners for [a], [ɛ] and [ɔ]. These modifications, however, are minimal. The data of Carignan (2018) and my own data suggest, that the vowels contribute to different degrees to the centralization of the vowel space due to the nasalization.

The results concerning the A1−P0 values reveal an adjustment in the amount of nasality only after exposure to high vowel nasalization. In this case, the amount of nasality in low vowels was decreased, whereas the amount of nasality in high and mid vowels was unaffected. As one would expect an increase of the amount of nasality after auditory exposure, this result might be due to the articulatory difficulty to nasalize high (and mid) vowels as proposed by ease of articulation. This failure to increase the amount of nasality in high (and mid) vowels was compensated by a decrease of the amount of nasality in low vowels, so that as a consequence the amount of nasality in high and mid vowels was indirectly increased.

Comparing the results of the experiment with the predictions, I can say that before exposure there was no difference in the amount of nasality across groups. This is in line with my prediction. I also assumed that all participants show the greatest amount of nasality in low vowels before exposure. This prediction is confirmed as well, however, in low vowels there was not a significantly greater amount of nasality than in the other vowels. Thus, participants showed a different pattern of nasalization in French loan words and in pseudowords: Whereas vowels in French loan words were all nasalized to the same degree, vowels in pseudowords were more heavily nasalized when they were low or mid than when they were high (see results of the production experiment in chapter 2.2.2). After exposure I expected a decrease of the amount of nasality with different predictions for the exposed vowels and for the unexposed vowels depending on exposure group.

When comparing the results of the experiment with the predictions concerning learning, I can say that the predictions of perceptual frequency and of ease of perception are confirmed: There was a greater increase of the amount of nasality in the exposed vowel after exposure to high vowel nasalization than after exposure to mid and low vowel nasalization. The prediction based on ease of articulation is not confirmed for learning. A summary of the predictions and the results can be seen in Table 12. In this Table green check marks represent confirmed predictions and red x marks represent unconfirmed predictions. Cells including '/ ' are those in which the phonetic effect makes no prediction. As the predictions based on perceptual frequency and ease of perception are identical, I cannot state whether the results are due to ease of perception or due to perceptual frequency.

Table 12: Predictions for the productive task in the artificial language learning experiment testing vowel nasalization and their evaluation.

	Learning		Generalizations of	
		[a] -learners	**[ɛ] -learners**	**[i] -learners**
Perceptual frequency	[i] > [ɛ], [a] ✓	/	/	/
Ease of perception	[i] > [ɛ], [a] ✓	high > mid X	high > low X	
Ease of articulation	[a] > [ɛ], [i] X		low > high X	low > mid X
Phonetic similarity	/	mid > high X	low > high X	mid > low ✓

A clear statement concerning the predictions for the generalizations is also difficult since only the participants of the [i]-group adjusted their amount of nasality significantly after exposure. The results of the participants of the [i]-group are in line with the prediction based on phonetic similarity. The prediction based on ease of articulation is again not confirmed. The predictions for the generalizations in the [a]- and [ɛ]-group are never confirmed because these participants did not generalize at all.

These conclusions are based on the results of the A1−P0 values which reached significance. The results of the centralization effect (vowel shifts and vowel space area) did not reach statistical significance but showed a tendency: They indicate a greater centralization effect (= more nasalization) after [ɛ]-exposure than after [i]-exposure, and a de-centralization effect (= less nasalization) after [a]-exposure due to the vowel space area. This is partly in line with the learning predictions based on perceptual frequency and ease of perception, but only for the de-centralization effect after [a]-exposure. Moreover, it can be observed in the data that after [a]-exposure there was the strongest shift of low vowels and after [ɛ]-exposure (and [i]-exposure) there was the strongest shift of high vowels. On the one hand, this shows that participants learned the nasalization pattern they were exposed to so that they increased the nasalization to the greatest degree in the exposed vowel height. On the other hand, it shows that the [ɛ]-group also increased their nasalization of high vowels which would be in line with the predictions based on ease of perception. In addition to that, there were more shifts to the central option after [a]-exposure than after [ɛ]-exposure and after [i]-exposure. This is in line with the learning predictions based on ease of articulation, which is contrary to the results based on A1−P0 values and calculations of the vowel space area.

To sum up, the predictions based on perceptual frequency, ease of perception and phonetic similarity are more often confirmed than the predictions based on ease of articulation. This is somewhat striking as the participants were asked to articulate nasal vowels. The previous artificial language learning experiment (see chapter 2.4) showed that all groups learned their nasalization pattern and that they were able to generalize this pattern to other patterns. As the exposure phase and the training phase were identical, I suggest that the participants in this production experiment also learned a nasalization pattern – either the nasalization of low, mid or high vowels. I expect that this learning can be observed in an increase of nasalization of this height for each exposure group. This is the case for the [i]-group, whose participants decreased the amount of nasality in low vowels based on A1−P0 values which indirectly increased the amount of nasality in high (and mid) vowels. This is the only significant result of the experiment. The calculations of the vowel space area as well as the choice of vowel options (see Table 10) are in line with this prediction.

Thus, the results show tendencies which illustrate that the participants learned their trained nasalization pattern as well – similar to the participants in the previous artificial language learning experiment. The results further show that participants exposed to high vowel nasalization increased the nasalization of their exposed vowel more than participants exposed to mid and low vowel nasalization. This may either be due to perceptual frequency or due to ease of perception. According to the prediction of perceptual frequency, participants exposed to nasal [ã] and nasal [ɛ̃] perceive nothing unfamiliar because they know these nasal vowels from French loan words. As a consequence they do not adapt their nasalization. The participants exposed to nasal [ĩ], however, perceive something unfamiliar – the nasalization of high vowels – which is not known from French loan words. As a consequence, they adapt their nasalization and want to increase the nasalization of their own high (and mid) vowels. However, as this is anatomically rather difficult for them, they decrease the nasalization of low vowels to indirectly increase the nasalization of high (and mid) nasal vowels. According to the prediction of ease of perception, participants know that oral and nasal high vowels are easiest to distinguish. Having perceived a nasal high vowel, they hence increase the nasalization of high vowels to a great degree. Having perceived a nasal low or mid vowel, however, they do not increase the nasalization that much so that the contrast between oral and nasal high vowels is still the easiest to perceive. When generalizing the nasalization pattern to other vowels no clear picture emerges. Whereas participants of the [i]-group generalized to mid vowels and participants of the [ɛ]-groups seem to have generalized to high vowels, participants of the [a]-group did not generalize at all. Thus, the generalizations may be based on phonetic similarity and on ease of perception. The inconclusive data for generalizations may either be due to several other factors which might have influenced this experiment, e.g. lexical frequency of loan words in German, unbalanced vowels in the dataset, individual differences and limited dataset, or due to the inability of German native speakers to produce high or mid nasal vowels appropriately so that it causes a significant increase or decrease of nasalization. More reliable data is thus necessary to resolve this issue.

3.4 Conclusion

The aim of this experiment was to find out whether auditory exposure to vowel nasalization of different height alters the production of the listeners – the degree of nasalization they produce – and if so, whether there is a difference in the degree of nasalization depending on whether they are producing a low, mid or high vowel in a nasal context. Doing so, I considered the phonetic effects of ease of articulation, ease of perception, phonetic similarity and perceptual frequency on the productions – similar to those effects on the perception in the artificial language learning experiment in chapter 2.4. I conducted an experiment in which three groups of German native speakers read aloud sentences with French loan words before and after auditory exposure to vowel nasalization. This auditory exposure included the presentation of singular, plural and diminutive forms in an artificial language created for the experiment described in chapter 2.4 in which plural was expressed by [ãm] ([a]-exposure), by [ɛ̃m] ([ɛ]-exposure) or by [ĩm] ([i]-exposure). The recordings were then analyzed to assess the degree of nasalization in the French loan words by measuring shifts in vowel quality, by measuring vowel spaces and by measuring acoustic nasalization based on $A1-P0$.

The results were highly variable so that none of the predictions could be entirely confirmed. The participants seem to have learned the nasalization pattern and consequently adjusted their production depending on the vowel height they were auditorily exposed to. Those participants who were exposed to high vowel nasalization increased the amount of nasalization most after exposure. This can either be due to the low perceptual frequency of high vowel nasalization in French loan words or due to perceptual ease. The results for the generalizations are inconclusive and suggest a bias towards phonetically similar patterns – if generalizations occurred at all. This is in line with the results of the previous artificial language learning experiment (see chapter 2.4) in which learning was based on perceptual ease and in which generalizations were based on phonetic similarity as well. Interestingly, the increase of nasalization is only achieved indirectly by decreasing the nasalization of dispreferred vowel heights as indicated by the acoustic measurement of nasality by means of $A1-P0$. The reason for this might be that German native speakers are apparently able to perceive nasalization properly, but have difficulties to articulate nasalized vowels or adjust the degree of nasalization depending on auditory exposure. The results of the [i]-group by means of $A1-P0$ suggest that German native speakers cannot increase the nasalization but have to decrease it in certain vowels to strengthen indirectly the nasalization of other vowels.

Although the results hardly reached significance, they, nevertheless, provide evidence for an effect of phonetics in phonological learning. This effect is stronger in perception than in production which is in line with the perception before production hypothesis (Flege, 1991) as a comparison of the results with the previous artificial language learning experiment which included a perceptual task shows. Moreover, it seems as if a generalization of perception onto production is difficult and that even a small modification of already known pronunciations depends on many factors, one of which is phonetics. In order to find out whether lexical effects confounded the results I suggest running the same experiment with pseudowords similar to those in the production experiment in chapter 2.2. This experiment may then result in greater changes in nasalization. I will leave that for future research.

4 Perception and learning of retroflexion

4.1 Introduction

How we perceive sounds and contrasts between them depends – among other factors, e.g. phonetic details, – on our native language (Abramson & Lisker, 1970; Best & Strange, 1992; Guion et al., 2000; Hallé, Best & Levitt, 1999; Kuhl & Iverson, 1995). Although all languages use the same sounds there are small phonetic differences among phoneme categories across languages, which make it difficult for speakers to perceive non-native sounds that sound similar to native ones. Japanese native speakers, for example, have difficulties differentiating the English liquid /l/-/r/-contrast because Japanese does not have this contrast. Whereas English has two phoneme categories – one for /l/ and one for /r/ – Japanese has only one phoneme category – /ɾ/. Consequently, Japanese native speakers perceive not only English /r/ but also English /l/ as Japanese /ɾ/ (Vance, 1987) (for details see chapter 1.1.3). Another example can be found in the perception of vowels: The English vowel contrast /iː/-/ɪ/ differs in terms of spectral and durational cues. Whereas native speakers of English mainly rely on spectral cues to differentiate these vowels, learners of English mainly rely on durational cues (Cebrian, 2006; Flege, Bohn & Jang, 1997; Rauber, Escudero, Bion & Baptista, 2005). Thus, phonetic details that are used to contrast sounds by native listeners may be unnecessary for L2-listeners, whereas L2-listeners may use phonetic details to contrast non-native sounds which are unnecessary for native listeners (Escudero, Benders & Lipski, 2009; Flege et al., 1997; Iverson et al., 2003). These perceptual difficulties have direct impacts on the ability to produce and to learn these sounds and contrasts (Aoyama et al., 2004; Best, 1995; Flege, 1995; Kuhl & Iverson, 1995; MacKay, Flege, Piske & Schirru, 2001).

In addition to the influence of the native language sound system, in this study I considered the effect of ease of perception on the learning behavior and on the generalization behavior of different learning groups. Doing so, I investigated the perception and learnability of retroflexion by German native speakers who are not familiar with retroflexion (Wiese, 1996). Retroflex consonants are those consonants which are articulated with a backward bended tongue tip (Trask, 1996). Acoustically they differ from alveolar consonants mainly in terms of a lowered F3 (Ladefoged & Maddieson, 1996; Narayanan & Kaun, 1999) (see chapter 1.4 for details). As fricatives are acoustically more salient than other consonant manners, a modification of fricatives, e.g. their retroflexion, causes a greater perceptual change than that of other consonant manners (Kohler, 1990). This is the reason why retroflex fricatives and alveolar fricatives are easier to distinguish than e.g. retroflex and alveolar plosives or retroflex and alveolar nasals.

Here, the effects of ease of articulation are not considered since they are inconclusive. Nevertheless, I will briefly explain which two different effects of ease of articulation are proposed – for the sake of completeness and to later show that neither of these articulatory effects affects learning and generalizing retroflexion. Whereas Stausland Johnsen (2012) mentioned no differences in the articulation among retroflex consonants, Hamann (2003b) pointed out that a fricative requires a longer retroflex gesture than plosives, nasals or lateral approximants. This longer duration might cause articulatory difficulties which result in articulatory ease for the retroflexion of plosives, nasals and lateral approximants compared to fricatives. However, neither of these two studies was an articulatory study and thus did not directly test the articulation of retroflexion. Stausland Johnsen (2012) studied the likelihood of retroflexion in Norwegian ([ʈ], [ɖ], [ɳ] > [ʂ]) and explained it well with the perceived distances between retroflex and alveolar consonants. Fricatives undergo less often retroflexion than plosives and nasals because the perceived distance between retroflex and alveolar fricatives is greater than the perceived distance between retroflex and alveolar plosives or nasals. This is in line with Steriade (2001b, 2009)'s P-map theory: The more perceptually distant x and x_1 are, the less likely does x surface as x_1. As Stausland Johnsen (2012) assumed no differences in the articulated retroflexion between consonant manners, articulation was not able to explain the likelihood of retroflexion in Norwegian. However, if there was an articulatory difference in the retroflexion of consonant manners, Stausland Johnsen (2012) would assume that retroflex fricatives are easier to articulate than retroflex plosives. He concluded this from the typological markedness of postalveolar (=retroflex) plosives which are less common than postalveolar (=retroflex) fricatives (Maddieson, 1984). Hamann (2003b) reviewed articulatory studies on retroflexion in different languages and grounded her assumption that retroflex fricatives are produced differently from retroflex plosives, nasals or lateral approximants on the articulatory differences in non-retroflex consonants. According to e.g. Dart (1991), Lindblad & Lundqvist (1999) and Wängler (1964), the fricative [s] is more laminal than the corresponding plosive [t]. The reason for this is that fricatives need a longer place of constriction – which is achieved by a more laminal articulation – compared to the small contact area of e.g. plosives to achieve the characteristic frication noise. That retroflex fricatives have indeed a different place of articulation and that they require a longer duration than the corresponding retroflex plosives was verified by e.g. Keating (1991), Ladefoged & Maddieson (1996) and Ladefoged & Wu (1984). As can be seen from this short comparison of two proposed effects of ease of articulation, it is not always clear how to define and to justify what is articulatory (or perceptually) easy. Consequently, learning biases based on ease of articulation (or ease of perception) as proposed by phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006) and their effects have to be evaluated carefully: It is not always clear what the theory predicts because it is not always clear which part of phonetics predicts which learning advantage and what part of phonetics is responsible for the typological distribution.

To assess the reported effects of ease of perception and the possible effects of ease of articulation – articulatory disadvantage of fricatives according to Hamann (2003b) and articulatory advantage for fricative or no articulatory advantage for any consonant manner according to Stausland Johnsen (2012) – I investigated the distribution of retroflexion in the languages of the world. Since ease of perception favors retroflex fricatives and ease of articulation may either favor retroflex plosives and retroflex lateral approximants, or retroflex fricatives or no retroflex consonant manner, I expected either an equal distribution of retroflex consonants among plosives, fricatives and lateral approximants (those manners I considered in the artificial language learning experiment) in the languages of the world or more retroflex fricatives than retroflex plosives or lateral approximants.

To explore the distribution of retroflexion across consonants in the languages of the world, I created a dataset of 91 languages with retroflex consonants from the UPSID corpus (Maddieson, 1984) as described in chapter 2.1. In this dataset, I annotated which retroflex manners of articulation were present for each language. Focusing only on the manners I examined in the perception experiments (plosives, fricatives, approximants, nasals) and in the learning experiment (plosives, fricatives, approximants), I can say that with 46% most of the languages that have retroflex consonants have retroflex plosives (n = 42). Retroflex approximants can be found in 26% (n = 24) and fricatives in 26% (n = 24) of the languages with retroflex consonants. Retroflex nasals are more rarely (15%, n = 14) (see Table 13).

However, these results can be misleading as they do not consider the number of occurrences of consonant manners without retroflex consonants in the world's languages. Thus, I calculated the percentages of retroflex consonants of different manners among the total number of consonants of different manners. 100% (n = 451) of all languages listed in the UPSID corpus (Maddieson, 1984) use plosives, 96% (n = 435) of all languages use nasals, 93% (n = 420) of all languages use fricatives and 96% (n = 434) of all languages use approximants (see Table 13). Calculating the percentages of retroflex consonants on these manners, it turned out that 9% of all plosives are retroflex, 6% of all approximants are retroflex, 6% of all fricatives are retroflex and 3% of all nasals are retroflex. Comparing the distribution of retroflex consonants with the distribution of retroflex consonants among all consonants (see Table 13), I can say that the tendency is the same with retroflex plosives being more frequent than retroflex fricatives and approximants which in turn are somewhat more frequent than retroflex nasals. However, based on the percentages the difference is smaller than in the raw numbers. These percentages show again (see typological distribution of nasal vowels in chapter 2.1) that the proposed encoding of the predispositions based on ease of perception and ease of articulation as markedness by the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006) is not met. Neither ease of perception (more retroflex fricatives than retroflex plosives and retroflex approximants) nor ease of articulation (more retroflex plosives and retroflex approximants than retroflex fricatives, or more retroflex fricatives than retroflex plosives and retroflex approximants, or no difference between consonant manners) nor a combination of both (equal distribution of retroflex plosives, retroflex fricatives and retroflex approximants, or more retroflex fricatives than retroflex plosives and retroflex approximants) can explain this typological asymmetry.

Table 13: Distribution of consonant manners in languages with consonants.

Consonant manner	**Retroflex C**	**C**	**Retroflex C among C**
plosive	42	451	9%
fricative	24	420	6%
approximant	24	434	6%
nasal	14	435	3%

In this study I thus excluded the possible effects of ease of articulation and typology and focused on perceptual effects only. I explored the role of perceptual distance in morphophonological learning. I focused on the perceptual distance between alternating sounds based on ease of perception and on the perceptual distance between native and non-native sounds – testing a possible interference of L1 with the latter. To do this I tested German native speakers on their perceptibility and learnability of retroflexion. Retroflex consonants are not part of the German phoneme inventory (Wiese, 1996) so that a direct influence of L1 can be excluded. However, there might be an indirect influence of the L1-sound system due to a perceptually similar sounding phoneme. First of all, I checked whether German native speakers can perceive retroflex consonants and whether there are differences depending on context and manner of articulation. I therefore conducted two perception experiments – one identification task and one discrimination task – with which I also established the perceptual distances between retroflex consonants and their alveolar counterparts (chapter 4.2-4.3). Next, I conducted an artificial language learning experiment to investigate whether there is a learning advantage for a specific manner of articulation depending on perceptual distances (chapter 4.4).

4.2 Perception experiment: identification task

Previous studies on the perception of retroflex consonants showed that the perception of retroflexion depends on three factors. The first factor is the consonant manner – with a perceptual advantage for fricative [ʂ] over plosive [ʈ] and [ɖ] and nasal [ɳ] (Stausland Johnsen, 2012). The second factor is the preceding vowel context – with a perceptual advantage for retroflex consonants following [a] over retroflex consonants following [i] (Ohala & Ohala, 2001) – and the third factor is the position of the retroflex depending on the vowel (VC vs. CV).

The results concerning the position are conflicting: Whereas Hamann (2003a) and Öhman (1966) found a perceptual advantage for retroflex consonants in VC over CV, Ahmed & Agrawal (1969) and Krull (1990) found a perceptual advantage for retroflex consonants in CV over VC. These conflicting results cannot be attributed to the native languages of the listeners since Krull (1990) and Öhman (1966) both tested Swedish native speakers. Moreover, the results cannot be explained by the consonants under investigation: Plosives were tested in all studies and in addition to this, Hamann (2003a) tested also nasals whereas Ahmed & Agrawal (1969) tested all Hindi consonants.

In order to investigate how speakers who are not familiar with retroflexion perceive the retroflexion of different consonant manners, I conducted a perception experiment in which I tested the perception of retroflexion with native speakers of German. I further aimed at investigating whether there is a perceptual advantage for retroflex consonants preceding vowels or for retroflex consonants following vowels. To this end, I tested the perception of retroflexion in both contexts – VC and CV. The vowel I chose was [a], as retroflexion was perceived best in this context in Ohala & Ohala (2001)'s study.

In this perception experiment the participants were asked to identify retroflex and alveolar consonants. This task tested the perceptual confusability of retroflex and alveolar consonants of different manners and in different positions (CV or VC). Based on the acoustics and Stausland Johnsen (2012)'s study, I predicted for the present experiment that if German native speakers benefit from a language-independent phonetic bias based on perceptual ease, they will confuse alveolar [s] and retroflex [ʂ] less often with each other than alveolar [t], [d], [n] and [l] with retroflex [ʈ], [ɖ], [ɳ] and [ɭ]. I added the lateral approximant [l] as a further test case and expect it to be perceived unlike the fricative and similar to plosives and nasals since the results of Stausland Johnsen (2012) suggest that retroflex fricatives are perceived better than all other retroflex consonant manners and as fricatives are acoustically more salient than other consonant manners (Kohler, 1990). If my participants do not confuse alveolar and retroflex consonants differently depending on consonant manner, they will not benefit from perceptual ease. In contrast to Ahmed & Agrawal (1969), Hamann (2003a), Krull (1990), and Öhman (1966) I investigated only the perception of retroflexion in speakers who are not familiar with retroflexion.

4.2.1 Method

Stimuli The stimuli were the five alveolar consonants [d], [t], [l], [n] and [s], which are part of the German phoneme inventory (Wiese, 1996), and their five retroflex counterparts [ɖ], [ʈ], [ɭ], [ɳ] and [ʂ], which are not part of the German phoneme inventory, either in [aC]-syllable or in [Ca]-syllable. The syllables were spliced out of [aCa]-syllables which were provided to me by a phonetic project of the University of Tromsø. The stimuli were recorded by a male monolingual native speaker of Norwegian. The sampling rate of the recording was 44.1 kHz. To avoid distracting noise at the beginning or at the end of the syllables, all syllables were sliced out of their context at the nearest zero-crossing point using Praat (Boersma & Weenink, 2019). As vowels show strong formants and a quasiperiodic wave form, I chose the beginning of the first quasiperiodic wave of the second vowel as endpoint of [aC]-syllables and I chose the end of the last quasiperiodic wave of the first vowel as beginning of [Ca]-syllables. This selection was in agreement with the visible formants in the spectrogram and my auditory impression. The intensity was then scaled to 70 dB. To make the task more difficult, babbling noise (55 dB) was added to the experimental items at a signal-to-noise ratio of about 7 dB using Praat (Boersma & Weenink, 2019). Babbling noise was also used by Stausland Johnsen (2012) to mask the stimuli. The babbling noise was the first part – *Jeg spiste...* – of the target sentence *Jeg spiste X til middag.* 'I ate X for dinner.' which was reversed. The target sentence was recorded by another male monolingual native speaker of Norwegian. Recording of the target sentence took place at the University of Tromsø. The sampling rate was 44.1 kHz. The practice consonants were not masked with noise.

Procedure The participants were tested with a forced-choice identification task using Praat MFC (multiple forced-choice) (Boersma & Weenink, 2019) on a Windows laptop. To avoid disturbing noise, they listened to the stimuli via headphones in a quiet room. In a short introductory phase the participants learned how the following stimuli sound and how they are represented orthographically on the screen as the participants were not familiar with retroflexion. They listened to each of the practice consonants once while they saw their IPA transcription on the screen. During the subsequent test the participants listened to a consonant masked with babbling noise and were forced to identify the consonant in each syllable as one of these: [d], [t], [l], [n], [s], [ɖ], [ʈ], [ɭ], [ɳ], [ʂ]. The participants responded by clicking on a box surrounding one of the consonants on the screen. After they had made their choice, the next syllable was presented after an inter-stimulus-interval of 500 ms. Each syllable was presented ten times in random order. The experiment lasted about ten minutes.

Participants I tested 20 adult native speakers of German (11 identified as female, 9 identified as male, mean age: 28.5, range: 19-54). No one reported knowledge of a language which uses retroflex consonants distinctively (except American English). All participants had normal or corrected-to-normal vision and no hearing problems. All of them participated voluntarily. Prior to the experiment each participant filled out a consent form and a questionnaire with which data about the participants' age, sex, native language, foreign language skills, the region they had been grown up and education was collected.

4.2.2 Results

The results of the perception experiment are provided in Tables 14 and 15. In these confusion matrices it is shown how often each stimulus (rows) was identified as one of the ten consonants (columns) in CV (see Table 14) and in VC (see Table 15). Percentages are given in brackets and shaded cells highlight the important comparisons. For example, the consonant [d] in the syllable [da] was identified as [d] 107 times – 89% of all responses given to the auditory stimulus [da] – and it was identified as [ɖ] 12 times – 10% of all responses given to the stimulus [da]. I was interested in how often each retroflex consonant was confused with its alveolar counterpart, and vice versa. The relevant cells are shaded in Table 14 and in Table 15.

The CV-results in Table 14 display an asymmetry. First of all, I can say that all alveolar consonants were easy to identify for German native speakers with percentages of correct responses between 78% and 93%. Second, the retroflex fricative [ʂ] was also quite easy to identify (73%). Third, the other retroflex consonants [ɭ], [ɳ], [ʈ] and [ɖ] were hard to identify with percentages of correct responses between 3% and 52%. All in all, the non-native retroflex consonants were more difficult to identify than the native alveolar consonants, except for retroflex [ʂ].

The VC-results in Table 15 are also asymmetric. First of all, the alveolar fricative [s] (85%) and the retroflex fricative [ʂ] (93%) were easy to identify. Second, the retroflex voiced consonants [ɳ], [ɭ], and [ɖ] were quite easy to identify with percentages of correct responses between 69% and 75%. However, the native alveolar voiced consonants [n], [l] and [d] were difficult to identify with percentages of correct responses between 47% and 53%. Further, they were often confused with their retroflex counterparts [ɳ], [ɭ] and [ɖ] with percentages between 42% and 51%. A possible reason might be that these consonants are articulated more front in Norwegian than in German (Kristoffersen, 2000) so that they sound different from what the listeners expected as their native alveolar consonants. Consequently, the participants were uncertain about what they heard and chose the non-native alternative. Third, the retroflex voiceless plosive [ʈ] and the alveolar voiceless plosive [t] were very hard to identify with percentages of correct responses of 27% and 36%. Further, they were not only confused with one another (8% and 31%) but also with the retroflex and alveolar voiced plosives [ɖ] and [d] (13%-52%). The same is true for retroflex voiced plosive [ɖ] which was sometimes identified as

Table 14: Confusion matrix of retroflex & alveolar consonants in CV-syllables in German.

stimulus \ response	[d]	[ɖ]	[t]	[ʈ]	[l]	[ɭ]	[n]	[ɳ]	[s]	[ʂ]
[da]	107 (89%)	12 (10%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	0 (0%)
[ɖa]	99 (83%)	20 (17%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	0 (0%)
[ta]	0 (0%)	0 (0%)	106 (88%)	13 (11%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)
[ʈa]	1 (1%)	0 (0%)	114 (95%)	3 (3%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)
[la]	0 (0%)	1 (1%)	0 (0%)	0 (0%)	93 (78%)	25 (21%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)
[ɭa]	0 (0%)	1 (1%)	0 (0%)	0 (0%)	56 (47%)	62 (52%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)
[na]	0 (0%)	1 (1%)	0 (0%)	0 (0%)	1 (1%)	1 (1%)	106 (88%)	10 (8%)	1 (1%)	0 (0%)
[ɳa]	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	64 (53%)	56 (47%)	0 (0%)	0 (0%)
[sa]	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)	1 (1%)	111 (93%)	7 (6%)
[ʂa]	0 (0%)	3 (3%)	0 (0%)	1 (1%)	1 (1%)	0 (0%)	1 (1%)	0 (0%)	26 (22%)	88 (73%)

Table 15: Confusion matrix of retroflex & alveolar consonants in VC-syllables in German.

stimulus \ response	[d]	[ɖ]	[t]	[ʈ]	[l]	[ɭ]	[n]	[ɳ]	[s]	[ʂ]
[ad]	60 (50%)	50 (42%)	2 (2%)	8 (7%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
[aɖ]	29 (24%)	83 (69%)	0 (0%)	7 (6%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)
[at]	21 (18%)	17 (14%)	43 (36%)	37 (31%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	1 (1%)
[aʈ]	15 (13%)	62 (52%)	9 (8%)	32 (27%)	1 (1%)	1 (1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
[al]	1 (1%)	0 (0%)	0 (0%)	1 (1%)	56 (47%)	61 (51%)	1 (1%)	0 (0%)	0 (0%)	0 (0%)
[aɭ]	0 (0%)	0 (0%)	0 (0%)	0 (0%)	33 (28%)	85 (71%)	2 (2%)	0 (0%)	0 (0%)	0 (0%)
[an]	0 (0%)	1 (1%)	0 (0%)	0 (0%)	1 (1%)	0 (0%)	63 (53%)	53 (44%)	0 (0%)	2 (2%)
[aɳ]	0 (0%)	1 (1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	29 (24%)	90 (75%)	0 (0%)	0 (0%)
[as]	1 (1%)	1 (1%)	1 (1%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)	102 (85%)	14 (12%)
[aʂ]	0 (0%)	1 (1%)	0 (0%)	1 (1%)	0 (0%)	0 (0%)	0 (0%)	1 (1%)	5 (4%)	112 (93%)

retroflex voiceless plosive [ʈ] (6%), and for the alveolar voiced plosive [d] which was sometimes identified as retroflex voiceless plosive [ʈ] (7%) or as alveolar voiceless plosive [t] (2%). Thus, not only retroflexion but also voicing seems to be difficult to identify in plosives.

All in all, the retroflex consonants were identified better in VC-context than in CV-context (see Tables 14 and 15). When ordering the retroflex consonants based on their percentages correct, however, both contexts did not differ: the voiceless fricative [ʂ] was easiest to perceive, the voiceless plosive [ʈ] was most difficult to perceive and the voiced nasal [ɳ], the voiced lateral approximant [ɭ] and the voiced plosive [ɖ] were somewhere in between. This is in line with the proposed ease of perception.

I calculated a linear mixed-effects model for each context separately on a subset of the data in R (R Core Team, 2015) with the corresponding R packages lme4 (Bates et al., 2015) and lmerTest (Kuznetsova et al., 2018). This subset included only the correct responses and those incorrect responses in which only retroflexion was misperceived. This corresponds to the shaded cells in Table 14 and in Table 15. In the models which fitted the data best (as assessed by backward stepwise elimination (Baayen, 2008)), ACCURACY (correct and incorrect) served as dependent variable and CONSONANT ([s], [l], [n], [d] and [t]) as independent variable. PARTICIPANT was included as random intercept (R syntax: glmer(ACCURACY ∼ CONSONANT + (1|PARTICIPANT), family = "binomial")).

The model for CV shows that alveolar and retroflex voiceless plosives as well as alveolar and retroflex voiced plosives were most often confused with each other. There was no significant difference in the confusions between alveolar and retroflex voiceless plosives and alveolar and retroflex voiced plosives (Est. $= -0.2964$, SE $= 0.1870$, z $= -1.585$, p $= 0.113$). Participants confused alveolar and retroflex lateral approximants significantly less often with each other than alveolar and retroflex voiceless plosives (Est. $= -0.8289$, SE $= 0.1923$, z $= -4.310$, p < 0.001) or alveolar and retroflex voiced plosives (Est. = -0.5325, SE = 0.1917, $z = -2.779$, $p < 0.01$). The same is true for confusions of alveolar and retroflex nasals: There were significantly less confusions of alveolar and retroflex nasals than of alveolar and retroflex voiceless plosives (Est. = -0.9600, SE = 0.1946, $z = -4.934$, $p < 0.001$) or of alveolar and retroflex voiced plosives (Est. = -0.6636, SE = 0.1939, $z = -3.422$, $p < 0.001$). There was no significant difference in the number of confusions between alveolar and retroflex lateral approximants and alveolar and retroflex nasals (Est. = 0.1310, SE = 0.1987, $z = 0.660$, $p = 0.510$). Alveolar and retroflex fricatives were least often confused with each other by the participants. The confusions of alveolar and retroflex fricatives differed significantly from all other consonants pairs ([t]-[ʈ]: Est. = -1.9959, SE = 0.2319, $z = -8.608$, $p < 0.001$; [d]-[ɖ]: Est. -1.6995, SE = 0.2312, $z = -7.351$, $p < 0.001$; [l]-[ɭ]: Est. = -1.1669, SE = 0.2349, $z = -4.968$, $p < 0.001$; [n]-[ɳ]: Est. = -1.0359, SE = 0.2368, $z = -4.375$, $p < 0.001$).

The results for VC are different. In this context alveolar and retroflex fricatives were confused with each other least often as well and the number of confusions differed significantly from all other consonant pairs ([t]-[ʈ]: Est. = -1.9471, SE = 0.3065, $z = -6.352$, $p < 0.001$; [d]-[ɖ]: Est. = -1.8427, SE = 0.2786, $z = -6.613$, $p < 0.001$; [l]-[ɭ]: Est. = -2.0382, SE = 0.2752, $z = -7.405$, $p < 0.001$; [n]-[ɳ]: Est. = -1.8157, SE = 0.2769, $z = -6.557$, $p < 0.001$). However, there was no significant difference in the confusions of all other consonant pairs. The number of confusions of alveolar and retroflex voiceless plosives did neither differ from that of alveolar and retroflex voiced plosives (Est. = 0.1044, SE = 0.2377, $z = 0.439$, $p = 0.660$), nor from that of alveolar and retroflex lateral approximants (Est. = -0.0910, SE = 0.2332, $z = -0.390$, $p = 0.696$), nor from that of alveolar and retroflex nasals (Est. = 0.1314, SE = 0.2353, $z = 0.559$, $p = 0.577$). Moreover, the number of confusions of alveolar and retroflex voiced plosives did not differ from those of alveolar and retroflex lateral approximants (Est. = -0.1955, SE = 0.1953, $z = -1.001$, $p = 0.317$) or alveolar and retroflex nasals (Est. = 0.0270, SE = 0.1978, $z = 0.137$, $p = 0.891$).

There was also no significant difference in the number of confusions between alveolar and retroflex nasals and alveolar and retroflex lateral approximants (Est. = −0.2224, SE = 0.1926, z = −1.155, p = 0.248).

In order to visualize the amount of confusion between alveolar and retroflex consonant pairs and to assess the perceived distance of retroflex and alveolar consonant pairs on the basis of the confusion matrices in Table 14 and 15, I calculated d'. I calculated d' using R (R syntax: `dprime.mAFC(proportion correct responses, number of alternative choices)`) and the corresponding R package psyphy (Knoblauch, 2014). The results can be seen in Figures 23 and 24 which were created with the package ggplot2 (Wickham, 2009) using R (R Core Team, 2015).[9]

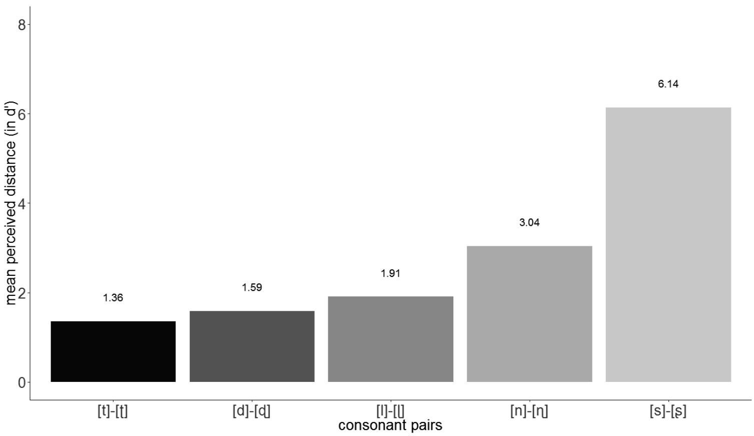

Figure 23: Perceived distance based on d' between retroflex and alveolar consonants in CV.

[9] d'-values can range from −10 to +10, whereby positive values mean that participants can discriminate the sounds above chance.

For CV-syllables (see Figure 23), the perceived distance based on d’ was largest between alveolar and retroflex fricatives with 6.14. The second largest distance was between alveolar and retroflex nasals with 3.04, followed by the distances between alveolar and retroflex lateral approximants (1.91), between alveolar and retroflex voiced plosives (1.59) and between alveolar and retroflex voiceless plosives (1.36).

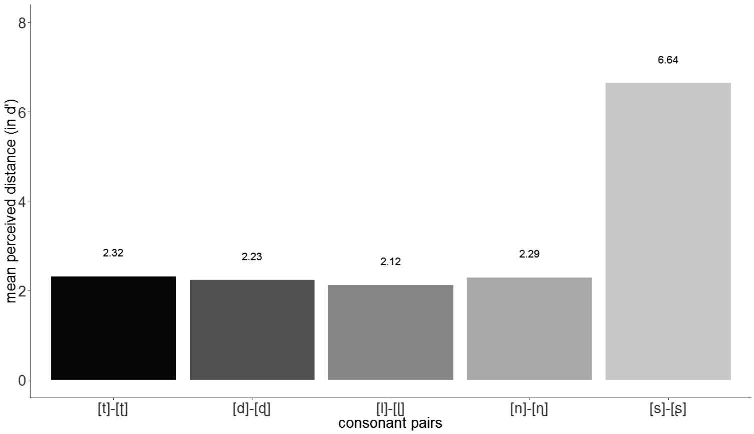

Figure 24: Perceived distance based on d’ between retroflex and alveolar consonants in VC.

For VC-syllables (see Figure 24), the perceived distance based on d’ was largest between alveolar and retroflex fricatives as well (6.64). The distances of the other pairs were very close to each other and varied from 2.32 between alveolar and retroflex voiceless plosives to 2.12 between alveolar and retroflex lateral approximants. The results for alveolar and retroflex voiceless plosives might be seen as contradictory to the data in the confusion matrix. In the confusion matrix it looks as if [ʈ] and [t] were very poorly identified, but the d’-results show no difference between [ʈ]-[t] and the other consonants pairs – except [ʂ]-[s]. However, the results are not contradictory: As d' is calculated based on the confusions of alveolar [t] with retroflex [ʈ], and vice versa, the confusions with retroflex [ɖ] and alveolar [d], which can be seen in Table 15, are not taken into account.

4.2.3 Discussion

I found two important results. First, retroflexion was perceived better in VC- than in CV-context and second, retroflex and alveolar fricatives were less often confused with each other than retroflex and alveolar plosives, nasals or lateral approximants.

The results confirm Hamann (2003a)'s and Öhman (1966)'s studies and show a perceptual advantage for retroflexion in VC- over CV-context – even for speakers who are not familiar with retroflexion. These results are also in line with Steriade (2001a) who stated that the perceptual distance in retroflex and alveolar consonant pairs is larger after a vowel than before a vowel – independent of consonant manner. An explanation why retroflex consonants are perceived better in VC than in CV is given by Hamann (2003b) and is conditioned by articulation. The characteristic movement for retroflex consonants is the backward curling of the tongue. This movement takes place at the beginning of the retroflex articulation. Thus, there is an articulatory difference in the transitions between vowels followed by a retroflex consonant and vowels followed by an alveolar consonant (VC). The articulatory movement at the end of the retroflex consonant does not differ that much from an alveolar consonant. Thus, there are no salient differences in the transitions between a retroflex consonant and an alveolar consonant into a vowel (CV) (Hamann, 2003b).

The results also confirm the pattern of Stausland Johnsen (2012) and show that [n], [l], [d] and [t] were more likely to be confused with [ɳ], [ɭ], [ɖ] and [ʈ] than [s] with [ʂ] – even by speakers who are not familiar with retroflexion. Stausland Johnsen (2012) did not give a clear explanation why the perceived distance between retroflex and alveolar fricatives is greater than that of the other consonant pairs, but related the perceptual results to the productivity of retroflexion in Norwegian – a language-specific effect – on the one hand, and to the greater salience of fricatives (Kohler, 1990) – a language-independent effect (ease of perception) – on the other hand. Ease of perception can also explain why there were identification advantages for fricatives independent of the context because this indicates a general perceptual advantage for retroflex fricatives – being present independent of experience with retroflexion. This is in line with Steriade (2001a)'s observations that the perceptual distance in retroflex and alveolar consonant pairs is greater for fricatives than for nasals and plosives. My data adds to this, showing that the perceptual distance between retroflex and alveolar lateral approximants is similar to that between retroflex and alveolar nasals and plosives.

However, it might also be possible that the similar results of Norwegian and German listeners are due to different language-specific reasons. Stausland Johnsen (2012) referred to the likelihood with which consonants undergo retroflexion in Norwegian. He found that the retroflexion of /t/, /d/ and /n/ in the context of /ɾ/ is obligatory, whereas the retroflexion of /s/ in the context of /ɾ/ is optional. He explained this observation with Steriade (2001b, 2009)'s P-map hypothesis: The smaller the perceived distance between retroflex and alveolar consonants is, e.g. between [t] and [ʈ], the more likely the alveolar consonant becomes retroflex. For German, I suggest that German native speakers map non-native retroflex [ʂ] onto the native postalveolar fricative [ʃ], which sounds very similar to the retroflex fricative and which is intermediate between alveolar [s] and retroflex [ʂ]. This mapping onto a native category does only happen with [ʂ] as there are no intermediate native categories for the other consonants as can be seen in Table 16. In this Table, a subset of the German consonant inventory and the retroflex consonants under investigation (in brackets) is displayed. There are no intermediate consonants between the retroflex and the alveolar consonants for plosives, nasals and lateral approximants. However, the postalveolar fricative is intermediate between retroflex fricative and alveolar fricative, so that retroflex fricatives might be mapped onto postalveolar fricatives.

Table 16: Subset of the German consonant inventory and retroflex consonants under investigation (in brackets).

	Alveolar	Postalveolar	Retroflex
Plosive	t d		(ʈ ɖ)
Nasal	n		(ɳ)
Fricative	s z	ʃ ʒ	(ʂ)
Lateral approximant	l		(ɭ)

As a consequence, [ʂ] might not be identified as retroflex but as postalveolar [ʃ] by the participants in the experiment. I cannot make any clear statement regarding fricatives concerning the perceptual advantage for retroflexion, as the participants might have failed to perceive the retroflexion in fricatives. If this is the case, there is no clear perceptual advantage for one of the retroflex consonants. To investigate this issue, I ran a second perception experiment (chapter 4.3).

4.3 Perception experiment: discrimination task

In an ABX perception experiment participants were asked to discriminate native German alveolar and postalveolar consonants and non-native retroflex consonants. This task mainly aimed at testing whether German native speakers can distinguish the non-native retroflex fricative [ʂ] from the native postalveolar fricative [ʃ]. However, I also wanted to test whether German native speakers are able to discriminate the non-native retroflex consonants and their native alveolar counterparts, or whether there is a mapping of retroflex consonants onto native categories independent of manner.

For the present experiment I predicted that native speakers of German are best at discriminating the native German contrast between [s] and [ʃ]. On the basis of the results of the previous perception experiment I further expected that German native speakers are able to discriminate between retroflex and alveolar fricatives, lateral approximants and plosives, whereby the contrast between [s] and [ʂ] is discriminated best and the contrast between [l] and [ɭ] is better discriminated than the one between [t] and [ʈ]. If German native speakers map the retroflex fricative onto the native postalveolar fricative, I will expect that they fail to discriminate the contrast between [ʂ] and [ʃ] and that the discrimination of the contrast between [s] and [ʃ] is the same as the discrimination of the contrast between [s] and [ʂ]. If German native speakers do not map the retroflex fricative onto the native postalveolar fricative, they will be able to discriminate the contrast between [ʂ] and [ʃ].

4.3.1 Method

Stimuli The stimuli were triplets (ABX) of [aC]-syllables. The consonants were either the non-native retroflex consonants [ɭ], [ʈ] and [ʂ] or the native (post)alveolar consonants [l], [t], [s] and [ʃ]. I excluded [d], [n], [ɖ] and [ɳ] as they were not used in the following artificial language learning experiment (chapter 4.4) due to similar patterning with [t], [l], [ʈ] and [ɭ]. The triplets always consisted of syllable pairs (A and B), which contained different consonants, and a third syllable (X), which was either another token of A or B. There were three tokens of each syllable in the experiment. The syllables which were compared in pairs were [l]-[ɭ], [t]-[ʈ], [s]-[ʂ], [ʃ]-[ʂ] and [s]-[ʃ]. These pairs always contained a non-native syllable (the one with the retroflex consonant) and a native syllable (the one with the alveolar consonant) – except for the pair [s]-[ʃ] which contained two native syllables.

There were a total of 60 triplets (12 triplets for each of the five comparisons). For each pair there was an equal number of triplets in which A and X contained the same consonant and in which B and X contained the same consonant. The order of presentation of the syllable pairs was balanced so that each of both syllables came first equally often.

The stimuli were recorded by a female fully bilingual Tamil-German native speaker who was able to produce the retroflex and (post)alveolar consonants accurately. Recording took place in an anechoic booth of the phonetics laboratory at Heinrich-Heine University Düsseldorf to avoid disturbing noise. For the recordings a microphone (phantom power 45 V), a Sound Devices amplifier, a Marantz recorder and a Transtec computer were used. The sampling rate was 44.1 kHz. Stimuli were recorded in the target sentence *Ich habe X oft gesagt.* 'I often said X.'. After recording the syllables were spliced out of the target sentence and their intensity was adjusted to 70 dB using Praat (Boersma & Weenink, 2019). The stimuli were always sliced at a zero-crossing point so that there were no irritations at the beginning or at the end of each stimulus.

Procedure The participants were tested with a forced-choice ABX task using Praat MFC (multiple forced-choice) (Boersma & Weenink, 2019) on a Windows laptop. In a quiet room they listened to the stimuli via headphones to avoid disturbing noise. Participants listened to each syllable triplet once in random order and were forced to identify which of the two first syllables (A or B) was repeated (X). The participants responded by clicking on a box with *1* or *2* on the screen. After they had made their choice the next triplet was presented.

After the presentation of all triplets one token of each syllable was randomly presented in isolation. The participants had to decide whether they know the consonant from the German language or whether they do not know it from German by clicking on a box with *Ja, kenne ich aus dem Deutschen.* 'Yes, it's a German sound.' or *Nein, kenne ich nicht aus dem Deutschen.* 'No, it's not a German sound.' on the screen. This was done to directly ask the participants whether the presented consonants – especially the retroflex fricative – are perceived as native. The experiment lasted about five minutes.

Participants I tested 30 adult native speakers of German (17 identified as female, 13 identified as male, mean age: 33.5, range: 18-62). No one reported knowledge of a language that uses retroflex consonants distinctively (except American English). All participants had normal or corrected-to-normal vision and no hearing problems. None participated in the previous perception experiment on retroflexion. All of them participated voluntarily. Prior to the experiment each participant filled out a consent form and a questionnaire with which data about the participants' age, sex, native language and foreign language skills was collected.

4.3.2 Results

For the ABX task 1800 data points (30 participants $\times$ 60 triplets) were analyzed in R (R Core Team, 2015) with the R package lme4 (Bates et al., 2015). I calculated the proportion of correct responses for each syllable pair separately. In the generalized linear mixed-effect model which fitted the data best (as assessed by backward stepwise elimination (Baayen, 2008)), ACCURACY (correct and incorrect) served as dependent variable and SYLLABLE PAIR ([s]-[ʃ], [s]-[ʂ], [t]-[ʈ], [l]-[ɭ] and [ʃ]-[ʂ]) as independent variable. PARTICIPANT and ITEM were included as random intercepts (R syntax: glmer(ACCURACY $\sim$ SYLLABLE PAIR + (1|PARTICIPANT)+ (1|ITEM), family = "binomial")). An overview of the results can be seen in Figure 25. This Figure was created using R (R Core Team, 2015) and the R packages Rmisc (Hope, 2013) and ggplot2 (Wickham, 2009).

Overall, the native contrast between the alveolar fricative [s] and the postalveolar fricative [ʃ] was discriminated best (86%), followed by the contrast between the alveolar fricative [s] and the retroflex fricative [ʂ] (85%). The percentage of correct responses for the contrast between the alveolar plosive

[t] and the retroflex plosive [ʈ] was 79% and for the contrast between the alveolar lateral approximant [l] and the retroflex lateral approximant [ɭ] it was 74%. The contrast between the postalveolar fricative [ʃ] and the retroflex fricative [ʂ] was discriminated worst (64%). All responses differed significantly from chance level ([s]-[ʃ]: Est. = 2.0043, SE = 0.2037, $z = 9.840$, $p < 0.001$, [s]-[ʂ]: Est. = 1.9143, SE = 0.2007, $z = 9.538$, $p < 0.001$, [t]-[ʈ]: Est. = 1.4611, SE = 0.1855, $z = 7.875$, $p < 0.001$, [l]-[ɭ]: Est. = 1.1418, SE = 0.1783, $z = 6.403$, $p < 0.001$, [ʃ]-[ʂ]: Est. = 0.6152, SE = 0.1710, $z = 3.598$, $p < 0.001$).

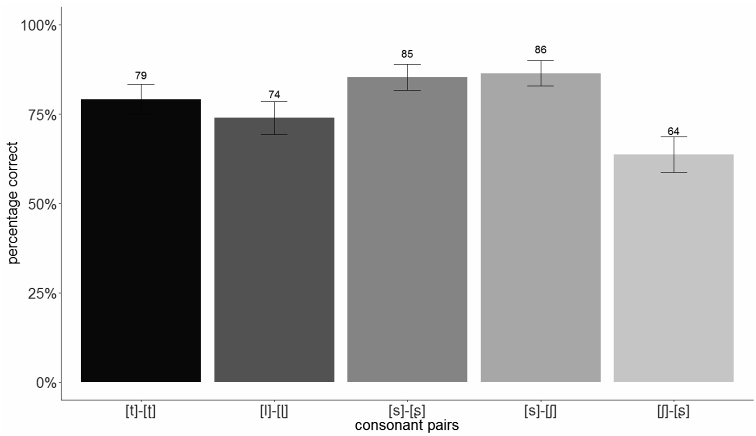

Figure 25: Percentages of correct responses ± 1.96 SE for each syllable pair in the discrimination task.

The percentage of correct responses to the native contrast [s]-[ʃ] did not differ significantly from those to the [s]-[ʂ] contrast (Est. = −0.0900, SE = 0.2228, $z = -0.404$, $p = 0.686$). There were, however, significantly more correct responses to the native contrast [s]-[ʃ] than to the [t]-[ʈ] contrast (Est. = −0.5432, SE = 0.2117, $z = -2.565$, $p < 0.05$), to the [l]-

[ɭ] contrast (Est. = −0.8624, SE = 0.2061, z = −4.185, p < 0.001) and to the [ʃ]-[ʂ] contrast (Est. = −1.3890, SE = 0.1989, z = −6.984, p < 0.001). Further, there were significantly more correct responses to the [s]-[ʂ] contrast than to the [t]-[ʈ] contrast (Est. = −0.4532, SE = 0.2078, z = −2.181, p < 0.05), to the [l]-[ɭ] contrast (Est. = −0.7725, SE = 0.2021, z = −3.822, p < 0.001) and to the [ʃ]-[ʂ] contrast (Est. = −1.2990, SE = 0.1941, z = −6.694, p < 0.001). There was no significant difference between the percentages of correct responses to the [t]-[ʈ] contrast and to the [l]-[ɭ] contrast (Est. = −0.3193, SE = 0.1881, z = −1.698, p = 0.895). However, there were significantly more correct responses to both contrasts than to the [ʃ]-[ʂ] contrast ([l]-[ɭ]: Est. = −0.5266, SE = 0.1750, z = −3.010, p < 0.01, [t]-[ʈ]: Est. = −0.8459, SE = 0.1820, z = −4.647, p < 0.001).

To answer the question whether the consonants are known from the German language or not, 210 data points (30 participants × 7 consonants) were analyzed. I calculated the percentages of correct responses using R (R Core Team, 2015). The results are shown in Table 17.

Table 17: Percentages of correct responses for the question whether the consonants are known from the German language or not.

Native consonant	**Percentage correct**	**Non-native consonant**	**Percentage correct**
[s]	83%	[ʂ]	67%
[ʃ]	70%		
[l]	87%	[ɭ]	80%
[t]	87%	[ʈ]	77%

German native speakers were most uncertain about the status of the non-native retroflex fricative [ʂ] – possibly because it sounds very much like the native postalveolar fricative [ʃ]. They were, however, also uncertain about the status of the native postalveolar fricative [ʃ] – possibly due to the task and the similar sounding retroflex fricative [ʂ]. The other German consonants [s], [l] and [t] were correctly identified as native with percentages of correct responses between 83% and 87%. The other non-native retroflex consonants [ɭ] and [ʈ] were correctly identified as non-native with percentages of correct responses between 77% and 80%.

4.3.3 Discussion

The main aim of this experiment was to investigate whether German native speakers can distinguish between the native postalveolar fricative [ʃ] and the non-native retroflex fricative [ʂ] or whether they do not perceive the retroflexion but map the non-native retroflex fricative [ʂ] onto the native postalveolar fricative category [ʃ]. The results show that the purely native contrast between the alveolar fricative [s] and the postalveolar fricative [ʃ] was perceived best. Interestingly, the results for the contrast between the alveolar fricative [s] and the retroflex fricative [ʂ] were the same, indicating that the postalveolar fricative [ʃ] and the retroflex fricative [ʂ] were perceived as similar in comparison to the alveolar fricative [s] by German native speakers. This is in line with the results for the contrast between the postalveolar fricative [ʃ] and the retroflex fricative [ʂ]: Only 64% of the participants were able to correctly perceive the contrast between the native postalveolar fricative and the non-native retroflex fricative. This was the worst result of all pairs. In addition to that, the retroflex fricative [ʂ] was the category with the worst results concerning the status as native or non-native consonant: 33% of the participants identified it as native consonant, probably as the postalveolar fricative [ʃ]. These findings altogether support my assumption that German native speakers map the retroflex fricative [ʂ] onto the native postalveolar fricative [ʃ].

The perceptual similarity of the alveolar fricative [ʂ] and the postalveolar fricative [ʃ] – which can be observed in the results of the German participants – may be the reason why there are only very few languages cross-linguistically that have both consonants in their inventories. As both consonants sound very similar, using both consonants contrastively might hamper communication. Consequently, this is expected to occur only very rarely, and this is exactly what I found: According to the UPSID corpus (Maddieson, 1984) only about 20 languages of a total of 451 languages have both fricatives – [ʂ] and [ʃ] – in their inventories. In addition to the rare contrastive use of retroflex and postalveolar fricatives, one has to keep in mind that only postalveolar fricatives exist. There are no postalveolar plosives, nasals or lateral approximants evidenced in the languages of the world according to the IPA chart. As such the contrastive use of retroflex and postalveolar consonants may be difficult due to other similar sounding consonants.

I was also interested in the discriminability of the contrast between retroflex and alveolar plosives and between retroflex and alveolar lateral approximants. In line with the VC-results of the first perception experiment (see chapter 4.2) both pairs did not differ from each other significantly, whereas both of these contrasts were more difficult to perceive than the contrast between retroflex and alveolar fricatives. The status as native or non-native category was correctly identified for all four sounds quite well (77%-87%).

The results of the present experiment show that both possible interpretations for the results of the first perception experiment (see chapter 4.2) might be true. I suggest a general perceptual advantage for retroflex fricatives based on ease of perception and an L1-specific perceptual advantage for retroflex fricatives for native German listeners. The latter assuming a mapping from the retroflex fricative [ʂ] onto the perceptually most similar native category [ʃ]. Based on the perception experiments alone I cannot state whether German native speakers rely on the language-general perceptual advantage or on the L1-specific perceptual advantage when listening to retroflex fricatives. However, when exposed to a retroflexion pattern in a learning experiment, participants should behave differently depending on what kind the perceptual advantage indeed is (language-general or L1-specific). Thus, I expect different generalization behaviors depending on whether learners perceive retroflex fricatives better than other consonants due to the acoustic salience of fricatives or due to a mapping onto the native category [ʃ]. As a mapping onto the native category [ʃ] means that no retroflexion pattern would be perceived in [s]-items, no generalizations from [s]-learners as well as no generalizations to [s]-items would confirm a language-specific advantage for retroflex fricatives. If there are, however, generalizations from [s]-learners as well as generalizations to [s]-items, the perceptual advantage for fricatives cannot be language-specific but tends to be language-general due to the acoustic salience of fricatives compared to other consonants. For details see chapter 4.4.

4.4 Artificial language learning experiment

The results of the previous perception experiments show that retroflex fricatives are generally perceived better than retroflex lateral approximants, and that retroflex plosives are generally perceived worst by German native speakers. In addition to that the results show that the perceptual distance between retroflex and alveolar consonants is generally greater in fricatives than in lateral approximants or plosives (especially in VC) for German native speakers and that German native speakers perceive retroflex fricatives as native postalveolar fricative and not as retroflex.

In order to test whether morphophonological learning is affected by the perceptual differences of retroflex and alveolar consonants of different manners – especially the perceptual distances between pairs of retroflex and alveolar consonants –, I conducted an artificial language learning experiment with German native speakers in which the crucial condition was the plural. The participants learned that plural forms ended in a retroflex consonant followed by [ɾa] (see phonological rule in (3)) and that diminutive forms ended in an alveolar consonant followed by [na]. The diminutives were added in order to conceal the focus of the experiment for the participants and to provide evidence that a consonant is only retroflexed in the plural. I chose the context VC[ɾ] because I found the best identification of retroflexion after vowels in the first perception experiment (chapter 4.2) and because the alveolar tap [ɾ] is the trigger for retroflexion in Norwegian according to Stausland Johnsen (2012). As in the artificial language learning experiment in chapter 2.4 and in the production experiment with French loan words in chapter 3, I made use of the poverty-of-the-stimulus design (Wilson, 2006): There were three groups of participants who all learned a retroflexion pattern but each with a different consonant manner – either with plosive [t], lateral approximant [l] or fricative [s]. In the subsequent test, however, the learners were asked to judge plural forms with consonants of all three manners.

(3) /s l t/ → [ʂ ɭ ʈ] / V_[ɾ]

There are no retroflex consonants in German (Wiese, 1996) but there is the postalveolar fricative [ʃ] in German which sounds similar to the retroflex fricative (see chapter 4.3.2). Because of this, I also tested whether phonological learning is affected more by language-specific perceptual characteristics or by language-general characteristics based on ease of perception. I chose the retroflex consonants [ʂ], [ɭ] and [ʈ] as they offer me the possibility to investigate the different predictions (see Table 18).

First, I lay out the predictions for learning – the participants' performance to the trained pattern. If there is a language-general learning bias that is based on ease of perception as proposed within the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004;

Wilson, 2006), I will expect that the retroflexion of fricatives is learned better than the retroflexion of lateral approximants and the retroflexion of plosives. The reason for this is that fricatives are acoustically more salient than other consonant manners so that modifications are perceived better in fricatives than in other consonant manners, e.g. plosives or lateral approximants. If there is a language-specific learning bias that is based on the perceptual similarity between non-native and native sounds, I will expect as well that the retroflexion of fricatives is learned better than the retroflexion of lateral approximants and the retroflexion of plosives. The reason for this is that the non-native retroflex fricative is perceived similar to the native postalveolar fricative so that there is a mapping of a non-native sound onto a native category. Thus, the fricative pattern is learned best because it represents a native contrast whereas the patterns with plosives and lateral approximants represent contrasts with non-native sounds.

Second, I lay out the predictions for the generalizations – the participants' performance to untrained patterns. These predictions differ depending on whether the perceptual advantage for fricatives is due to ease of perception or due to a mapping onto a native category. If there is a learning bias that is based on ease of perception, I will expect that [l]-, [s]- and [t]-learners all generalize more to [s]-items than to [l]- or [t]-items. If ease of perception is part of their linguistic knowledge, they know that alveolar and retroflex fricatives are easiest to distinguish. Having perceived a retroflex consonant of any manner (and they can do so in principle, see Table 15), they nevertheless have a preference to generalize to fricatives. If there is a learning bias that is based on the perceptual similarity to native language sounds, however, I will expect that [s]-learners do not generalize at all. This is because [s]-learners do not perceive and learn a retroflexion pattern but an alternation between the alveolar fricative and the postalveolar fricative – an alternation which cannot be found in the retroflexion patterns concerning the alternation of alveolar and retroflex plosives and lateral approximants. Learners from the [l]-group are expected to generalize their retroflexion pattern more to [t]-items than to [s]-items because they perceive the retroflexion in [t]-items but not in [s]-items due to the mapping of the retroflex fricative onto the postalveolar fricative. Finally, [t]-learners are expected to generalize more to [l]-items than to [s]-items because they perceive the retroflexion in [l]-items but not in [s]-items due to the mapping of the retroflex fricative onto the postalveolar fricative.

Table 18: Predictions for the artificial language learning experiment testing retroflexion.

	Learning	**Generalizations of**		
		[t] -learners	**[s] -learners**	**[l] -learners**
Language-general perception	[s] > [l], [t]	[s] > [l]		[s] > [t]
	[s] > [l] > [t]	–		[s] > [t]
Language-specific perception	[s] > [l], [t]	[l] > [s]	–	[t] > [s]
	[s] > [l] > [t]	–	–	[t] > [s]

Table 18 provides an overview of the predictions for the artificial language learning experiment testing retroflexion. Empty cells are those in which the consonant in training would be the one most generalized to. In this case the effect of phonetics and the effect of exposure during the training phase of the experiment would not be separable. In the second lines the predictions for the case that [ʈ] is perceived as [ɖ] are shown. Due to the perceptual confusions of retroflex [ʈ] with retroflex [ɖ] as shown by the results of the first perception experiment (see Table 15) it might be the case that [t]-learners do not only learn a retroflexion pattern but also perceive an alternation of voicing. If this is the case, [t]-learners are expected to perform worst during learning as they have to learn a more complex pattern which is interpreted as an alternation of two features (place of articulation and voicing). They are further expected to fail to generalize their learned pattern to other consonants. This is because neither the contrast between [l] and [ɭ] nor the contrast between [s] and [ʂ] or [s] and [ʃ] involves a voicing contrast.

4.4.1 Method

Stimuli The artificial language consisted of singular, plural and diminutive forms. The stimuli were constructed from a subset of the German (Wiese, 1996) and Norwegian phoneme inventories (Kristoffersen, 2000). All items were in agreement with the phonotactics of German (Wiese, 1996) – except for the retroflex consonants and [ɾ]. The structure of the stimuli and examples of each grammatical form are illustrated in Table 19. The singulars were pseudowords of the form $C_1V_1C_2$ with [s], [l] or [t] as C_2. The plural was expressed by the final syllable [ɾa], which caused the retroflexion of the preceding consonant, and the diminutive was expressed by the final syllable [na] without any other phonological change. The consonants [p], [b], [d], [k], [g], [f] and [v] were used in C_1-position. In V_1-position I used the vowels [ɛ], [ɪ], [ɔ] and [ʊ].

Table 19: Structure of the stimuli for the artificial language learning experiment testing retroflexion.

Form	**C_1**	**V_1**	**C_2**	**Suffix**	**Example**
singular	[p b d k g f v]	[ɛ ɪ ɔ ʊ]	[s l t]	∅	[bɛl]
plural	[p b d k g f v]	[ɛ ɪ ɔ ʊ]	[ʂ ɭ ʈ]	[ɾa]	[bɛɭɾa]
diminutive	[p b d k g f v]	[ɛ ɪ ɔ ʊ]	[s l t]	[na]	[bɛlna]

For the training phase I used 48 items for each group of participants (16 singulars, 16 plurals and 16 diminutives). For the subsequent test phase I used 48 stimulus pairs, each consisting of a form which conformed to the retroflexion rule (e.g. correct plural [bɛɭɾa] and correct diminutive [bɛlna]) and one which did not (e.g. incorrect plural [bɛlɾa] and incorrect diminutive [bɛɭna]). Half of the pairs (n = 24) tested the plural formation and half of the pairs (n = 24) tested the diminutive formation. In both stimulus groups there was an equal number of eight pairs for each of the three consonants. Half of these pairs were part of the training items (n = 4) and half of them were not (n = 4). These numbers are the same as in the artificial language learning experiment on vowel nasalization (see chapter 2.4.1). The complete set of all stimuli used in the artificial language learning experiment on retroflexion is listed in the appendix (see Tables 31, 32, 33, 34).

The stimuli were recorded by a male monolingual native speaker of Norwegian. Recording took place at the University of Tromsø. The sampling rate was 44.1 kHz. Stimuli were recorded in the target sentence *Jeg spiste X til middag.* 'I ate X for dinner.' to focus on the stimulus item and to ensure a uniform language environment that allows the speaker to naturally produce retroflex consonants. To avoid irritating noise at the beginning or at the end of the items, all stimuli were sliced out of their context at the nearest zero-crossing point after recording using Praat (Boersma & Weenink, 2019). The intensity of all stimuli was adjusted to 70 dB using Praat (Boersma & Weenink, 2019).

Procedure The procedure was the same as in the artificial language learning experiment testing vowel nasalization (see chapter 2.4) – except that the retroflexion experiment ran on a Mac computer. For details see the description of the procedure in chapter 2.4.1.

The experiment was divided into a perceptual training phase and a perceptual forced-choice test. It was scripted in PsychoPy (Peirce, 2007) and it ran on a Mac computer. Participants listened to the auditory stimuli via headphones. The experiment lasted about 10 minutes and took place in an anechoic booth in the phonetics laboratory of the Heinrich-Heine University Düsseldorf to avoid disturbing noise.

Participants were trained on a subset of the stimuli but tested on all stimuli – including stimuli known from training and stimuli they had not yet heard. There were three training groups: During training each group heard the retroflexion of only one consonant – either [s] or [l] or [t]. For example, a member of the [s]-group never received any training for [l] or [t]. The test was identical for all groups and included all consonants.

At the beginning of the experiment participants were informed that a new language was discovered on an island in the South Sea and that they would get to know this new language as well as the animals living on this island in the course of the experiment.

In a short introductory phase the participants were familiarized with the set-up of the experiment by listening to German animal names in the singular, plural and diminutive. The auditory stimuli were supported visually with images (Snodgrass & Vanderwart, 1980) to illustrate their meaning. The introductory phase was followed by the training phase and the test phase.

In the training phase the participants listened to two repetitions of 48 stimuli (16 singulars, 16 plurals and 16 diminutives) in randomized order. An auditory stimulus was played while an image of a fantasy animal (van de Vijver & Baer-Henney, 2014) was shown (see Figure 12). A singular form was accompanied by a single fantasy animal, a plural form by two fantasy animals and a diminutive form by a small fantasy animal. During training participants received positive input only, which means that they never listened to an incorrect plural, e.g. [bɛlɾa], or to an incorrect diminutive, e.g. [bɛɭna].

The test following the training phase was a perceptual forced-choice task and it was the same for all groups. There were 48 stimulus pairs consisting of a correct and an incorrect form differing only in the retroflexion of C_2. Half of the pairs ($n = 24$) tested the plural formation and half of the pairs ($n = 24$) tested the diminutive formation. There were 16 pairs with [s], [l] and [t]. A trial consisted of the presentation of the first form, followed by the presentation of the second form (see Figure 13). There was an equal number of trials in which the first and the second form was the correct one. The auditory stimuli were supported visually either by an image of two fantasy animals (plurals) or by an image of one small fantasy animal (diminutives). The participants had to decide which of the two forms was the correct one by either pressing the right or the left arrow key within 3000 ms. All stimulus pairs were presented in randomized order.

Participants 68 native speakers of German took part in the experiment (46 identified as female, 21 identified as male, 1 identified as gender diverse, mean age: 26.2, range: 19-63). They were randomly assigned to one of the three experimental groups. 20 participants were trained with retroflexion of [l], 26 participants were trained with retroflexion of [s] and 22 participants were trained with retroflexion of [t]. All groups were tested with all consonants [l], [s] and [t]. No one reported knowledge of a language that uses retroflex consonants distinctively (except American English). All participants had normal or corrected-to-normal vision and no hearing problems. None participated in the previous perception experiments on retroflexion (identification task and discrimination task). Participants were given a small expense allowance for their participation. Prior to the experiment each participant filled out a consent form and a questionnaire with which data about the participants' age, sex, native language and foreign language skills was collected.

4.4.2 Results

From 3264 data points (68 participants $\times$ 48 trials) 38 were not analyzable because the participants did not respond within 3000 ms. I analyzed the proportion of correct answers from the remaining 3226 data points with R (R Core Team, 2015) by means of a generalized linear mixed-effects analysis with the corresponding R package lme4 (Bates et al., 2015).

I was interested in the learning and generalization behavior across experimental groups. An overview about the results provides Figure 26. This Figure was created using R (R Core

Team, 2015) and the R packages Rmisc (Hope, 2013) and ggplot2 (Wickham, 2009). First of all, I compared the results of the trained consonant in each group. This is what I call *learning*. In the model which fitted the data best (as assessed by backward stepwise elimination (Baayen, 2008)), ACCURACY (correct and incorrect) served as dependent variable and TRAINED CONSONANT ([l], [s] and [t]) was the independent variable. PARTICIPANT and ITEM were included as random intercepts (R syntax: glmer(ACCURACY $\sim$ TRAINED CONSONANT + (1|PARTICIPANT) + (1|ITEM), family = "binomial")).

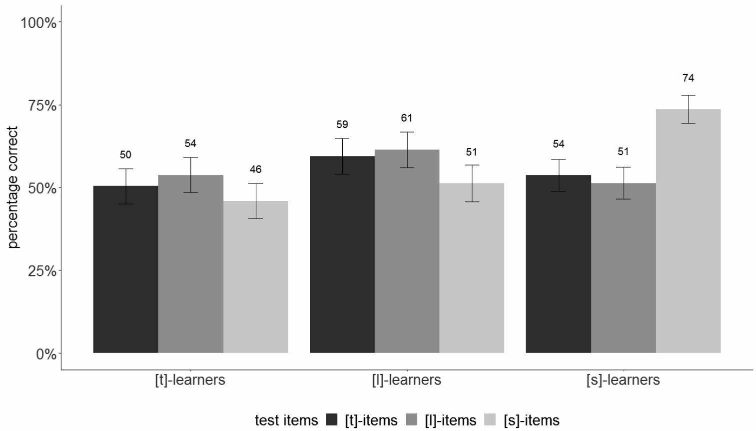

Figure 26: Percentages of correct responses ± 1.96 SE across all training groups and all consonant manners at test (1).

In the trained condition the percentage of correct responses of [s]-learners was 74%, that of [l]-learners was 61%, and that of [t]-learners was 50%. The percentage of correct responses of [l]-learners and [t]-learners did not differ from each other (Est. = −0.4916, SE = 0.3113, z = −1.579, p = 0.114). Both groups gave significantly less correct responses to items in their trained condition than [s]-learners to items with [s] ([l]-learners: Est. $= -0.6722$, SE $= 0.3165$, $z = -2.124$, $p < 0.05$; [t]-learners: Est. $= -1.1638$, SE $= 0.3137$, $z = -3.710$, $p < 0.001$). The percentage of correct responses to the trained items differed from chance level for [s]-learners (Est. $= 1.227$, SE $= 0.2584$, $z = 4.749$, $p < 0.001$) and for [l]-learners (Est. $= 0.5012$, SE $= 0.1841$, $z = 2.722$, $p < 0.001$) but not for [t]-learners (Est. $= 0.0226$, SE $= 0.1981$, $z = 0.114$, $p = 0.909$).

I also investigated how participants judged items with untrained consonants. This is what I call *generalization*. Therefore I performed three generalized linear mixed-effects models – one for each learning group. In the models which fitted the data best (as assessed by backward stepwise elimination (Baayen, 2008)), ACCURACY (correct and incorrect) served as dependent variable and CONSONANT AT TEST ([l], [s] and [t]) as independent variable. PARTICIPANT and ITEM were included as random intercepts (R syntax: glmer(ACCURACY ~ CONSONANT AT TEST + (1|PARTICIPANT) + (1|ITEM), family = "binomial")).

[t]-Learners responded correctly to [t]-items in 50%, to [l]-items in 54% and to [s]-items in 46%. Neither the percentages of correct responses to [s]-items (Est. $= -0.1819$, SE $= 0.1998$, $z = -0.910$, $p = 0.363$) nor those to [l]-items (Est. $= 0.1754$, SE $= 0.2000$, $z = 0.877$, $p = 0.380$) differed significantly from chance level. The percentage of correct responses to [t]-items did not differ significantly from that to [s]-items (Est. $= -0.2044$, SE $= 0.2814$, $z = -0.726$, $p = 0.468$), nor from that to [l]-items (Est. $= 0.1529$, SE $= 0.2815$, $z = 0.543$, $p = 0.587$). The percentages of correct responses to [l]-items and to [s]-items did not differ significantly from each other either (Est. $= -0.3573$, SE $= 0.2827$, $z = -1.264$, $p = 0.206$).

[l]-Learners responded correctly to [l]-items in 61%, to [t]-items in 59% and to [s]-items in 51%. The percentages of correct responses to [t]-items differed significantly from chance level (Est. $= 0.4084$, SE $= 0.1831$, $z = 2.231$, $p < 0.05$), those to [s]-items did not (Est. $= 0.0547$, SE $= 0.1817$, $z = 0.301$, $p = 0.764$). The percentage of correct responses to [l]-items

differed marginally significantly from that to [s]-items (Est. = −0.4465, SE = 0.2557, z = −1.747, p = 0.081). There was no significant difference between the percentages of correct responses to [l]-items and to [t]-items (Est. = −0.0928, SE = 0.2564, z = −0.362, p = 0.717). The correct responses to [s]-items and to [t]-items differed significantly from each other (Est. = −1.0565, SE = 0.3497, z = −3.022, p < 0.01).

[s]-Learners responded correctly to [s]-items 74%, to [t]-items in 54% and to [l]-items in 51%. Neither the percentage of correct responses to [l]-items (Est. = 0.0565, SE = 0.2479, z = 0.228, p = 0.820) nor that to [t]-items (Est. = 0.1707, SE = 0.2472, z = 0.691, p = 0.490) differed significantly from chance level. The percentage of correct responses to [s]-items differed significantly from that to [t]-items (Est. = −1.0565, SE = 0.3497, z = −3.022, p < 0.01) and from that to [l]-items (Est. = −1.1708, SE = 0.3503, z = −3.342, p < 0.001). There was no significant difference between the percentages of correct responses to [t]-items and to [l]-items (Est. = −0.1142, SE = 0.3422, z = −0.334, p = 0.738).

4.4.3 Discussion

The artificial language learning experiment shows that native speakers of German are able to learn a retroflexion pattern. Their performance differed regarding consonant manner: Participants trained with fricatives learned their trained pattern best, followed by participants trained with lateral approximants. Participants trained with plosives did not learn their retroflexion pattern as their percentages of correct responses did not differ from chance level. This learning advantage for fricatives is in line with the expectation based on ease of perception which states that the greater the perceived distance between two sounds is – in this case between alveolar and retroflex consonants –, the easier it is to learn the contrast between these two sounds (Aoyama et al., 2004; Best, 1995; Flege, 1995; Kuhl & Iverson, 1995). The reason for this is that the perceptual distance between fricatives is acoustically more salient than the perceptual distance between other consonant manners (Kohler, 1990). However, these results are also in line with the expectations based on a mapping of non-native sounds onto native sounds. As proposed by Aoyama et al. (2004), a non-native sound which is very similar to a native sound is learned better than a non-native sound which is dissimilar to a native sound. This is because the non-native sound is directly mapped onto an already existing native sound so that no new sound category has to be established. Participants trained on retroflex fricatives perceived the non-native retroflex fricative as native postalveolar fricative so that they perceived and learned a pattern involving only native sounds, whereas participants trained on retroflex plosives or retroflex lateral approximants perceived and learned a pattern involving native and non-native sounds. Since learning a familiar contrast between two native sounds is easier than learning an unfamiliar contrast, participants trained on fricatives had a familiarity advantage. Whereas Aoyama et al. (2004) proposed that a non-native sound which is perceptually most similar to a native one is learned easiest, Flege (1995) and Kuhl & Iverson (1995) claimed that a perceptually more distant non-native sound is learned easier than one that is perceptually close to a native sound. However, as in the case of Aoyama et al. (2004)'s proposal the non-native sound is actually mislearned as a native category both statements do not contradict each other. The learning results further show that participants exposed to retroflex plosive [ʈ] performed worst. The reason for this might be that they perceived retroflex [ʈ] as retroflex [ɖ]. This is a wrong mapping, too, however, it is no mapping onto a native category. Thus, this mapping does not result in a learning advantage. It is rather a learning disadvantage as the contrast between the alveolar voiceless plosive [t] and the retroflex voiced plosive [ɖ] involves a two-feature change. [ʈ]-Learners therefore suffered from a complexity bias (Moreton, 2008, 2012).

Whereas the predictions for the learning results are the same for the language-general perception account and for the language-specific perception account, which means that both of them are confirmed, the predictions for the generalization results differ depending on the account (see Table 20).

The [s]-learners were most successful in items with their trained consonant [s]. However, they did not generalize their learned retroflexion rule to items with the untrained consonants [l] and [t]. Their percentage of correct responses to items with lateral approximants and plosives was almost identical and near chance level. This result is in line with the language-specific perception account. The participants did not perceive the retroflex fricative as retroflex but mapped the non-native sound onto the native postalveolar fricative [ʃ]. At test, [s]-learners were also confronted with pairs of retroflex and alveolar plosives and lateral approximants, but not with postalveolar plosives or lateral approximants – both of which do not exist at all. Thus, the participants were not able to generalize their learned pattern which involves – according to their perception – an alternation between alveolar and postalveolar consonants.

The [l]-learners were successful in items with their trained consonant [l] as well. In contrast to the [s]-learners, the [l]-learners generalized their retroflexion pattern to items with the untrained consonant [t]. However, they did not generalize it to items with the untrained consonant [s]. This result is again in line with the language-specific perception account but not with the language-general perception account. Participants perceived and learned a pattern involving alveolar and retroflex lateral approximants. At test, [l]-learners were confronted with alveolar and retroflex plosives, to which they generalized their learned pattern. They were, however, also confronted with alveolar and postalveolar fricatives – [l]-learners did not perceive the retroflexion in fricatives either. This did not conform to the pattern they know from training so that they were not able to generalize their learned pattern to fricatives.

The [t]-learners, did not learn retroflexion at all. As a consequence they were not able to generalize a retroflexion pattern to items with the untrained consonants [s] and [l]. They made no difference between items with trained and items with untrained consonants and performed at chance level in all three conditions. Interestingly, [t]-learners gave more – but not significantly more – correct responses to [l]-items than to [s]-items. This latter tendency would be in line with the language-specific perception account but not with the language-general perception account. Thus, it might be the case, that [t]-learners perceived one part of their pattern (retroflexion, but not voicing) in [l]-items but not in [s]-items (in which they neither perceived a retroflexion nor a voicing alternation). This would be a further hint at the language-specific perception account. However, these differences did not reach significance, so that I suggest that [t]-learners did not learn the pattern they were exposed to.

To sum up, the results of the learning experiment can be explained by the influence of L1-perception. The participants seem to have perceived retroflex [ʂ] as the native postalveolar fricative [ʃ]. This suggestion is in line with the results of the second perception experiment (see chapter 4.3). As a consequence, participants trained with [s]-items perceived (and learned) a contrast between two native sounds – [s] and [ʃ]. During training they did not perceive any retroflexion pattern at all. In the subsequent test, they were confronted with items with retroflex plosives and retroflex lateral approximants. These retroflex sounds were completely new for them as they had not perceived retroflexion before. The percentages of correct responses to items with plosives and lateral approximants are thus at chance. Participants trained with [l]-items perceived retroflexion of [ɭ] in training and generalized it to [t]-items, but not to [s]-items at test. The reason for this might be that the retroflexion of [ʈ] at test was perceived so that the [l]-learners were able to apply their learned retroflexion pattern not only to [l]-items but also to [t]-items. However, [l]-learners did not perceive [ʂ] as retroflex [ʂ] but as postalveolar [ʃ] – similar to the [s]-learners. Thus, the [s]-items did not include any retroflexion pattern, which they had learned before. Consequently, their performance with [s]-items was at chance. Interestingly, the native sound categories seem to have a very strong 'magnet' effect: Even though [l]-learners had already perceived retroflexion in the trained items, they did not perceive the fricative as retroflex but as postalveolar. Hence, the familiarity with native sound categories has a stronger influence than the familiarity with a non-native sound category having been exposed to only some minutes ago. Participants trained with [t]-items had the most difficult task. As they did not perceive and learn the retroflexion of [t] at all, they were not able to generalize it to other consonants. In the results of the [t]-learners I observe the most correct responses to [l]-items, followed by the responses to [t]-items and to [s]-items. Although this difference was not significant, it shows once more the disadvantage in generalizing to fricatives.

An overview of the predictions and the experimental results can be found in Table 20. In this Table green check marks represent confirmed predictions and red x marks represent unconfirmed predictions. In the second lines the predictions for the case that [ʈ] is perceived as [ɖ] are shown. Empty cells are those in which the consonant in training would be the one most generalized to. In this case the effect of phonetics and the effect of exposure during the training phase of the experiment would not be separable.

Comparing the results of the learning experiment with the predictions I can say that the predictions for learning are met not only by the language-specific perception account but also by the language-general perception account – when considering the second line of each prediction as [t]-learners seem to have perceived retroflex [ʈ] as [ɖ]. In learning, those contrasts which contain the greatest perceptual distance are learned best. In this case it does not matter onto which kind of category (native or non-native) both sounds are mapped – what counts is that both sounds are perceptually far away from each other and thus mapped onto two different categories. Those different categories can either be both native, both non-native or one native and one non-native – in all three cases the contrast is learned. Note, however, that this is only true for alternations which are of equal complexity. Alternations which involve a more complex change, e.g. [t]-[ɖ], do not profit from a great perceptual distance as structural complexity hampers learning them.

Table 20: Predictions for the artificial language learning experiment testing retroflexion and their evaluation.

	Learning	**Generalizations of**		
		[t] -learners	**[s] -learners**	**[l] -learners**
Language-general perception	[s] > [l], [t] ✗ [s] > [l] > [t] ✓	[s] > [l] ✗ – ✓		[s] > [t] ✗ [s] > [t] ✗
Language-specific perception	[s] > [l], [t] ✗ [s] > [l] > [t] ✓	[l] > [s] ✗ – ✓	– ✓ – ✓	[t] > [s] ✓ [t] > [s] ✓

The predictions for the generalizations of the [s]-learners are met by the language-specific perception account. [s]-learners did not learn retroflexion at all and were thus not able to generalize this pattern to items showing a retroflex contrast. The predictions for the generalizations of the [l]-learners are only met by the language-specific perception account and not by the language-general perception account. [l]-Learners generalized their retroflexion only to [t]-items because they did not perceive the retroflexion of [ʂ]. According to the language-general account the opposite would be true, meaning that [l]-learners perceived the retroflexion of [ʂ] and generalized it to [s]-items but not to the perceptually more similar contrast in [t]-items. The predictions for the generalizations of the [t]-learners are met by both accounts when assuming that retroflex [ʈ] was perceived as retroflex [ɖ] as in the first perception experiment (see Table 15). If this is the case, [t]-learners had the most difficult task in identifying the alternation as retroflexion (and not as voicing alternation). This difficulty may have led to a wrong rule (similar to that of the [s]-learners). Consequently, this wrong rule could not be applied to the contrastive stimulus pairs at test which did not include a voicing contrast. [t]-Learners were thus not able to generalize at all – similar to the [s]-learners.

I therefore conclude that the language-specific perception account can best explain the results. Not only the learning results but also all generalization results can be explained with assuming that non-native retroflex [ʂ] was perceived as native postalveolar [ʃ] by the participants.

Summing up, I found that a contrast of two sounds which are perceptually more distant from each other is learned easier than a contrast of two sounds which are perceptually more close to each other. Moreover, I found a dependency on the perceptual distance to native sound categories: Whereas learning is enhanced by a direct mapping from a non-native category onto a native one – caused by a very small perceptual distance between non-native and native category –, generalizing this learned pattern to other consonants is hampered by a direct mapping onto a native category – probably because the learners did not perceive and learn what they were supposed to perceive and learn[10].

4.5 General discussion

The aim of this study was to investigate the role of perceptual distances to native and non-native sounds in morphophonological learning. I looked at the perceptual distances between alternating sounds and at the perceptual distances between non-native alternating sounds and perceptually similar native sounds onto which the non-native sound may be mapped, to test whether language-general perceptual effects as instantiated by

[10] A possible phonetic bias grounded in articulatory ease would have resulted either in a learning and generalization advantage for retroflex fricatives (Stausland Johnsen, 2012) with the same predictions as those based on language-general perception, in a learning and generalization advantage for neither consonant manner (Stausland Johnsen, 2012) or in a learning and generalization disadvantage for retroflex fricatives (Hamann, 2003b), neither of which are confirmed by the experimental results.

ease of perception or language-specific perceptual effects as instantiated by a mapping onto native sound categories influence the learning behavior and the generalization behavior, and, if so, whether one of these effects overwrites the other one.

To this end I investigated the perception and learnability of retroflexion of different consonant manners by native speakers of German who have no knowledge of retroflexion in three experiments. The first perception experiment showed that the perception of different consonant manners differed. Whereas retroflex fricatives were identified best, retroflex plosives were identified worst. Retroflex lateral approximants were somewhere in between. The perceived similarity between retroflex and alveolar consonants differed as well. Retroflex and alveolar fricatives were perceptually more distant from each other than retroflex and alveolar lateral approximants or plosives. These perceptual results are similar to the ones by native speakers of Norwegian who are familiar with retroflexion (Stausland Johnsen, 2012). It is thus possible that there is a general perceptual advantage for retroflex fricatives over other retroflex consonants. However, the second perception experiment also showed that the perceptual distance between alveolar and retroflex fricatives and between alveolar and postalveolar fricatives were the same and that retroflex fricatives and postalveolar fricatives were perceptually very close to each other – suggesting a possible mapping of retroflex fricatives onto native postalveolar fricatives by native speakers of German. This means that the perceptual advantage for retroflex fricatives over other retroflex consonants may be due to the perceptual proximity to a similar sounding L1-sound.

I conducted an artificial language learning experiment in which three groups of German native speakers were taught a language in which plural was expressed by the retroflexion of an alveolar consonant – either of a fricative ([s]-learners) or of a plosive ([t]-learners) or of a lateral approximant ([l]-learners) – to test whether morphophonological learning is affected more by a language-general perceptual advantage or by a language-specific perceptual advantage. Therefore, I subsequently tested the participants on their generalizations to all consonants. The results showed that perceptual distances to native and non-native sounds influence morphophonological learning and how the learned pattern is generalized to other consonant manners. Participants learned the retroflexion of fricatives better than the retroflexion of other consonant manners. This finding shows that the perceptual distance influences morphophonological learning: The more perceptually distant two sounds are, the better this contrast is learned. This is in agreement with Aoyama et al. (2004), Flege (1995) and Kuhl & Iverson (1995).

During learning it does not matter whether one (or both) of the involved sounds of a contrast is incorrectly mapped onto a native category. What counts is just the perceptual distance between the perceived sounds. However, when generalizing the learned contrast to other sounds, it is important how the involved sounds are mapped. My participants have mapped the retroflex fricative onto the native postalveolar fricative. The group of [s]-learners thus did not learn any retroflexion and did not generalize its learned rule ('make alveolar consonants postalveolar') to the other consonants as the presented items at test did not allow for such a rule. Thus, although being the best group in learning, they were not able to generalize – due to an L1-effect. The group of [l]-learners learned retroflexion but generalized it only to [t]-items, not to [s]-items as they perceived the retroflex fricatives as postalveolar – and not as retroflex – as well. The [t]-learners had the most difficult task in figuring out what the rule is, as the identification of retroflex plosives was the most difficult for German native speakers. Unfortunately, they failed to learn their pattern at all and were thus unable to generalize retroflexion to other consonants. If they had learned their pattern, I would have expected generalizations to [l]-items but not to [s]-items, which would be in accordance with the language-specific perception account.

This study does not only shed light on the different effects of perceptual distances to native and non-native sound categories in learning and generalizations, but also argues for the controversial role of phonetics in synchronic phonology. As Kingston (2007) (page 401) put it "the field has reached no consensus about what the interface is, nor has it even agreed that one exists at all". While the role of phonetics was argued to be purely diachronic by some researchers (Blevins, 2004; Ohala, 1993; Yu, 2004), others explained their synchronic experimental results as a consequence of a phonetic bias (Baer-Henney & van de Vijver, 2012; Finley, 2009, 2012; van de Vijver & Baer-Henney, 2014; White, 2017; White & Sundara, 2014; Wilson, 2006) and argued that there is a phonetic influence on synchronic phonology (Hayes & Steriade, 2004).

On the one hand, Blevins (2004) argued against a learning bias and assumed that phonetics shapes phonology diachronically exclusively. Difficulties in perception result in re-analysis and cause changes in phonology over time, which leads to typological asymmetries. However, they cannot be used synchronically by language learners. When two sounds are acoustically and perceptually similar, they are often confused which each other. As a consequence, the contrastive use of these sounds hampers effective communication. To overcome this difficulty, the contrast between these sounds disappears gradually over time. The use of sounds which are perceptually more distant from each other encounters no such perceptual problem. In the case of retroflexion this would mean that the retroflex and alveolar fricatives can co-exist in the languages of the world because they are perceptually distant from each other, whereas retroflex and alveolar plosives cannot because they are perceptually close to each other. My typological survey, however, did not prove this (see chapter 2.1). Moreover, this would mean that postalveolar and retroflex fricatives cannot co-exist in one language. It was rather difficult to prove this but it seems to be the case: I searched in the literature for languages which have both fricatives, especially to find an appropriate speaker for the stimuli in the ABX task (chapter 4.3), but I found only very few languages which have both sounds (28 languages, based on combined data from UPSID (Maddieson, 1984) and PHOIBLE (Moran & McCloy, 2019)). In addition to that, in personal communications with speakers of languages which seem to have retroflex fricatives, e.g. Norwegian, I learned that native speakers are not sure of whether the fricative is retroflex or postalveolar. Even Stausland Johnsen (2012) referred in his paper on retroflexion to postalveolar consonants. This shows that a process may currently taking place by which one of both sounds is replaced by the other one.

On the other hand, proponents of the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006) argued for a role of phonetics in synchronic phonology and highlight the diversity of phonetic effects on phonology, e.g. phonological processes are motivated by perceptual or articulatory cues. In the phonological process of place assimilation neighboring segments agree in their places of articulation which is determined by articulatory phonetics (Jun, 2004) and the phonological distribution of contour tones depends on the duration and the sonority of the syllable rhyme as contour tones need a long and sonorous rhyme to be properly articulated and perceived (Zhang, 2004). In addition to that distinctive features are mainly defined in terms of phonetic properties but are used to explain phonological processes and even phonological representations can be described in terms of phonetic properties (Kingston, 2007). Furthermore, there are many experimental studies, e.g. Baer-Henney & van de Vijver (2012), Finley (2009, 2012), van de Vijver & Baer-Henney (2014), White (2017), White & Sundara (2014) and Wilson (2006), which show, similar to the experiments I present here, that phonetics influences phonology in the form of a phonetic bias during learning. For example, Finley (2012) studied the learnability of vowel harmony and found that a harmony triggered by mid vowels was easier to learn than a harmony triggered by high vowels, which can be explained with perceptual ease (Kaun, 2004). Thus, small phonetic details – in this case perceptual details – affect phonological learning. However, some of the studies mentioned above confounded a phonetic bias with structural simplicity (Moreton & Pater, 2012a,b; Pater & Moreton, 2012). I avoided this confound by using retroflexion of alveolar consonants of different manners. The alternation from alveolar to retroflex is structurally akin for every manner. However, there are perceptual differences for each manner. This allowed me to provide more solid evidence for a role of phonetics in learning synchronic phonology. The same is true for the learning experiment on vowel nasalization (see chapter 2.4).

The misperceptions of the participants could have resulted in a representation of different contrasts than the one intended. As mentioned before, retroflex voiceless plosives were identified by more than half of the native German participants as retroflex voiced plosives in VC-context (see Table 15). This means that [t]-learners were perceptually confronted with a two-feature change between alveolar voiceless plosives and retroflex voiced plosives. The other learning groups, however, were only confronted with a one-feature change from alveolar to retroflex (or postalveolar). Thus, the [t]-learners' failure to learn their rule cannot only be explained by uncertainty about what exactly the pattern is but also by a learning disadvantage due to a complexity bias (Moreton, 2008, 2012).

Results of artificial language learning experiments are further often interpreted in the light of typology, e.g. studies by Finley (2009), White (2013) and Wilson (2006). However, the results of my learners do not mirror the typological distribution of retroflex consonants and it is even not the case that the phonetic factor ease of perception (language-general perception) matches the typological distribution of retroflex consonants (see chapter 4.1). According to the typological distribution of retroflexion, I would expect a learning advantage and a generalization advantage for plosives over lateral approximants and fricatives. I found the opposite result – an advantage for fricatives and a disadvantage for plosives. The results of the artificial language learning experiment thus show – similar to those of the artificial language learning experiment on vowel nasalization (see chapter 2.4) – that the phonetic knowledge about phonology is not necessarily encoded as markedness and reflected typologically. Independent of this, ease of perception – and also ease of articulation – may affect the phonology of languages diachronically. As sound patterns in languages were – and still are – affected by many factors, e.g. phonetics, not all learning biases directly affect typological distributions.

Diachronically, there seems to be evidence for a perceptual advantage for retroflex fricatives as well. In Indic retroflex /ʂ/ seems to have been developed from /ʃ/. According to Blevins (2017) and Hamp (1996) the reason for this shift is language contact with the Dravidian language in which the retroflex sibilant /ʂ/ was already present as well as bilingualism at the time of contact. Due to the perceptual magnet effect (Kuhl, 1991) the phonetically closest Indic sound shifted to the retroflex category – *ʃ became retroflex /ʂ/ (Hamp, 1996). Blevins (2017) further suggested that the Dravidian speakers might have mapped their sibilant contrast between /s/ and retroflex /ʂ/ onto the Indic sibilant contrast between /s/ and /ʃ/. Indic plosives, laterals and nasals became retroflex due to assimilation processes with retroflex /ʂ/ (Hamp, 1996). This assimilation process is partly in agreement with the language-general perceptual account based on ease of perception: Retroflexion of /s/ is learned best and generalized to the same degree to other consonant manners. However, as the data of the artificial language learning experiment shows, the shift from /ʃ/ to /ʂ/ is very unlikely, even in a retroflexion facilitating context before /ɾ/. Due to the results of the experimental study, it is thus more likely that retroflexion emerged from the fusion of alveolar consonants with /ɾ/, as Sarmā (2004) suggested e.g. for retroflex plosives in Ladakhi.

4.6 Conclusion

The present study investigated the role of perceptual distances in morphophonological learning. Perceptual distances can be measured between alternating sounds and between native and non-native sounds when testing a non-native alternation. In general, an alternation is learned better if the alternating sounds are perceptually distant from each other, whereas an alternation is poorly learned if the alternating sounds are perceptually close to each other (Aoyama et al., 2004; Flege, 1995; Kuhl & Iverson, 1995). I tested native speakers of German who were not familiar with retroflexion to investigate which effects perceptual distances have not only on learning a phonological alternation but also on generalizing the learned alternation to other consonants. I showed that German native speakers identified retroflex fricatives best, followed by retroflex lateral approximants and retroflex plosives (see chapter 4.2). The same is true for the perception of the contrast between retroflex and alveolar consonants (see chapter 4.3): The contrast between retroflex and alveolar fricatives was discriminated best, followed by the contrast between retroflex and alveolar lateral approximants and the one between retroflex and alveolar plosives. I further showed that German native speakers discriminated the contrast between retroflex and postalveolar fricatives only poorly – suggesting a mapping of non-native retroflex fricatives onto native postalveolar fricatives. In an artificial language learning experiment I found that the retroflexion of fricatives was learned easier than the retroflexion of other consonants (see chapter 4.4). This mirrors the differences in the perceptual distances between alternating sounds and a mapping onto a perceptually close native fricative. However, participants trained with fricatives did not generalize at all and participants trained with other consonants did not generalize to fricatives. This means that a mapping onto a native fricative category enhances learning, but hampers when it comes to generalizations. The reason for this is that a mapping onto a native category – in this case – means a mapping onto a wrong category in terms of the alternation that should be learned. Generalizing the actually learned alternation was thus impossible. The same was true for true retroflexion learners who were not able to apply the retroflexion rule to the perceived item pairs of alveolar and postalveolar fricatives. I thus provided evidence for a direct influence of perception and an even stronger influence of L1 in morphophonological learning. This influence of L1 is not always helpful but can also be hampering.

5 Learning morphophonology without alternation

5.1 Introduction

In the previous artificial language learning experiments (see chapters 2.4, 3.2 and 4.4) learners were always confronted with a morphophonological alternation. They learned that the singular takes a specific form and that the plural is formed by adding a suffix causing a phonological alternation, e.g. nasalization of the oral singular vowel in the plural or retroflexion of the alveolar singular consonant in the plural, whereas the diminutive is formed by adding a suffix to the singular form without any further phonological alternation. Learners were thus confronted with three forms of the noun paradigm in the nominative case: nominative singular, nominative plural and nominative diminutive. This was done to provide the learners with detailed information about the morphophonological pattern in the artificial languages. As the results of the artificial language learning experiments have shown, this information was used by the participants to learn how the plural and the diminutive are formed in order to be able to generalize this formation to novel items. However, did the participants really learn an alternation (X $\rightarrow$ Y, e.g. oral vowel in the singular $\rightarrow$ nasal vowel in the plural or alveolar consonant in the singular $\rightarrow$ retroflex consonant in the plural) or did they just learn that the singular and the diminutive contain sound X (e.g. an oral vowel or an alveolar consonant), whereas the plural contains sound Y (e.g. a nasal vowel or a retroflex consonant) – independent of any alternation? That is, did the participants learn a relation between forms in a paradigm or did they transfer e.g. the feature [nasal] from a known item to a novel item?

To investigate these questions I re-ran the artificial language learning experiment on retroflexion (see chapter 4.4) with a short modification: Participants were not confronted with singular forms, but with plural and diminutive forms only. I decided to omit singulars as singulars are generally considered to be the base forms in German which do not contain a phonological alternation, whereas plurals or diminutives may contain a phonological alternation of the singular form (Eisenberg & Sayatz, 2004; Köpcke, 1988; Wurzel, 1965). Consider, for example, the plural and diminutive formation in German. Not only in plurals but also in diminutives vowel and voicing alternations are possible (Wiese, 1996): The singular form of 'mouse' is *Maus* [maʊs], the plural form is *Mäuse* [mɔʏzə] and the diminutive form is *Mäuschen* [mɔʏsçən]. In this case, the vowel and the fricative in the singular are the underlying sounds which have a different surface form in the plural and in the diminutive than in the singular due to a vowel alternation ([s]-[z]) and a voicing alternation ([aʊ]-[ɔʏ]). When learners are exposed to singulars, plurals and diminutives, it is not possible to answer the abovementioned questions and to decide which of the two strategies the learners made use of. However, when learners are exposed to plurals and diminutives only – meaning the lack of the base form singular – one can investigate whether learners need a relation between forms to accurately acquire how the plural and the diminutive are formed or whether they are able to transfer the feature of a given form to unknown items without a relation between forms. Whereas the first option would imply an alternation between sound X and sound Y, the second option does not. Thus, if participants are able to accurately learn and generalize the retroflexion pattern in the artificial language learning experiment without being exposed to singulars, participants will not learn an alternation but the phonetic and phonotactic properties of plurals and diminutives. However, if participants are not able to accurately learn and generalize the retroflexion pattern in the artificial language learning experiment without being exposed to singulars, all forms of a paradigm – or at least the base form singular – need to be presented to accurately predict the phonological forms and their relation among each other.

Support for the idea of transfer (without alternation) comes from Berent (2013) and Marcus (2001). They proposed that learners form phonological generalizations on the basis of algebraic rules (Berent, Marcus, Shimron & Gafos, 2002; Berent et al., 2012). By using these rules learners generalize not only to novel items with native sounds (Berent, Everett & Shimron, 2001; Berent, Vaknin & Marcus, 2007) but even to novel items with non-native sounds, and as such to features that are unattested in the native language of the learners (Berent et al., 2002, 2012). In the case of the learning experiment on retroflexion, this means on the one hand that learners should be able to learn that plural forms contain the feature [retroflex] – independent of the presence or absence of this feature in other forms – and on the other hand that learners should be able to generalize the feature [retroflex] e.g. from plural forms with a plosive to plural forms with a lateral approximant or a fricative, and vice versa. Note that lateral approximants and fricatives are non-native sounds for the [t]-learners, as they did not listen to [l]- and [s]-items during training (see chapter 4.4 and 5.2). Berent (2013)'s and Marcus (2001)'s assumption is based on experiments with Hebrew speakers (Berent et al., 2001, 2002, 2007). Whereas Hebrew does not allow two identical consonants at the left edge of the word (*AAB), e.g. *[sisum], two identical consonants at the right edge of the word (ABB) are allowed, e.g. [simum], (Greenberg, 1950). In lexical decision tasks Hebrew speakers were confronted with AAB-items and with ABB-items. The consonants in the items were either consonants which were part of the Hebrew phoneme inventory or consonant which were not part of the Hebrew phoneme inventory. The results showed that the participants classified AAB-items faster and more often as non-words than ABB-items. This was true for items with native consonants (Berent et al., 2001, 2007) as well as for items with non-native consonants (Berent et al., 2002). Thus, Hebrew speakers were able to recognize phonotactic constraints and the presence of certain phonetic features in the presented items and to transfer (or generalize) them to novel items. Consequently, exposing learners to plurals and diminutives only – and not to the singular bases – can shed light on the question whether learners can transfer one phonological form to another one without knowledge about a relation between forms of a paradigm or whether learners need a relation between forms of a paradigm to learn the presented morphophonological pattern.

The predictions for the outcome of the artificial language learning experiment depending on the learning strategy are shown in Table 21. I termed the learning strategies *transfer account* and *relation account*. If participants rely on the transfer account, I will expect them to accurately learn and generalize the retroflexion pattern without being exposed to singular forms. Participants can do so, as they will not learn an alternation but the phonetic and phonotactic properties of plurals and diminutives. These properties will then be transferred from the trained items to the untrained items to the same degree. If participants rely on the relation account, I will expect them to not accurately learn and generalize the retroflexion pattern without being exposed to singulars. Participants cannot do so, as they need to know the base form singular to accurately predict the phonological forms and their relation among each other.

Table 21: Predictions for the artificial language learning experiment testing retroflexion without alternation (1).

		Generalizations of		
	Learning	**[t] -learners**	**[s] -learners**	**[l] -learners**
Transfer account	✓	✓	✓	✓
Relation account	✓	–	–	–

In addition to the information about the learning strategies, the lack of singular forms can also shed further light on the interaction of language-specific (= L1) and language-general perceptual cues (= phonetics) as will be elaborated below.

The previous artificial language learning experiment (see chapter 4.4) showed that participants rely more on language-specific perceptual cues than on language-general perceptual cues based on ease of perception and that L1 can be helpful in learning but also hampering when something is mislearned due to a wrong mapping on a native category. To my knowledge there is only one study which directly aimed at testing the interaction of language-general and language-specific factors during learning. Baer-Henney et al. (2015a) investigated the role of phonetic substance and of L1-phonotactics on the learnability of phonological patterns. Native German participants were either exposed to a vowel harmony pattern which was phonetically motivated but not present in the German language or to an arbitrary vowel alternation pattern which was not phonetically motivated but conformed to German phonotactics (Wiese, 1996). In addition to that each pattern contained exceptions, which means that either 85% or 65% of all training items conformed to the pattern being exposed to. One further factor under investigation was the length of training which included either two or three repetitions of all training items. The results showed that participants who had received a long training performed better at test than those who had received a short training and that participants who had been exposed to fewer exceptions performed better at test than those who had been exposed to more exceptions. In general, participants exposed to the arbitrary pattern with few exceptions performed best – that are those participants who had been exposed to a pattern which conformed to their L1-phonotactics. This means that L1-phonotactics seems to have a greater influence on phonological learning than phonetics. Only when exposed to a training set with many exceptions to the pattern and with a short training phase, participants relied on phonetics. The longer the training phase and the fewer exceptions, the more strongly were the results influenced by L1-phonotactics. The effect of length of training and amount of exposure is in line with Baer-Henney & van de Vijver (2012): Training sets with 50% singulars and

50% alternating plurals led to better performances at test than training sets with 75% singulars and 25% alternating plurals.

Similar results could be observed in two experiments of Greenwood (2016). In her artificial language learning experiments native speakers of English were trained either on a phonetically natural pattern or on a phonetically unnatural pattern. In addition to that, each pattern was presented either in casual speech or in careful speech, whereas casual speech corresponds to natural speech and careful speech to a clear (hyperarticulated) speech. The experiment described in chapter 6 in Greenwood (2016) studied the learnability of a phonetically unnatural final voicing pattern and a phonetically natural final devoicing pattern with fricatives. The results showed that in careful speech participants performed better in the phonetically unnatural condition than in the phonetically natural condition. However, in casual speech there was no difference in performance between the phonetically unnatural and the phonetically natural condition. Thus, phonetics did not seem to have any effect on the results in careful speech, as the results did not display a learning advantage for the phonetically natural pattern. However, as there are more final voiced fricatives than final voiceless fricatives in English, participants may have been influenced by their L1-lexicon in careful speech, which led to the better performance in the final voicing condition (Greenwood, 2016). This influence of the L1-lexicon could not be found in casual speech which I interpret as a more difficult and less reliable training condition. Thus, the less reliable training in casual speech downgraded the influence of L1 so that participants performed equally well with phonetically natural and phonetically unnatural patterns. Note that the performance in the phonetically unnatural condition decreased from careful speech (77%) to casual speech (63%) but that the performance in the phonetically natural condition was equal in careful speech (57%) and casual speech (59%). Greenwood (2016)'s experiment described in chapter 7 studied the learnability of a phonetically unnatural phonotactic pattern which allows only obstruents as codas and a phonetically natural phonotactic pattern which allows only sonorants as coda. In careful speech – the more reliable training condition – there was no difference in performance between the phonetically natural and the phonetically unnatural pattern. In casual speech – the less reliable training condition –, however, participants performed better with the phonetically natural pattern than with the phonetically unnatural pattern. This shows that under less reliable conditions participants' performance is more affected by phonetics than under more reliable conditions, which is in line with the results of Baer-Henney et al. (2015a).

Baer-Henney et al. (2015a)'s and Greenwood (2016)'s results may explain why the artificial language learning experiment on retroflexion in chapter 4.4 did not reveal a clear influence of a phonetic bias based on perceptual ease in phonological learning but an influence of perceptually similar L1-sounds. Participants were exposed to 100% conforming items during training and they always listened to singulars, plurals and diminutives. The latter being added to give evidence for an alternation which occurs only in plurals. The same is true for the former artificial language learning experiments on vowel nasalization (see chapters 2.4 and 3). However, an influence of L1 could not show up as there are no perceptually similar sounds for the nasal vowels in German – except for the corresponding oral vowels, which would result in a non-alternation, e.g. if the nasal vowel [ĩ] in the alternating pattern [i]-[ĩ] was mapped onto the oral vowel [i], it would results in the non-alternating pattern [i]-[i].

However, confronting participants with a more difficult task may result in a phonetic bias emerging in learning retroflexion, which would provide evidence for the language-general perception account based on ease of perception, and in less influence of the L1. The task in the present experiment would be more difficult than in the previous artificial language learning experiment on retroflexion when e.g. singulars would not be presented so that evidence for an alternation would be missing and when the training phase would in general be shorter. The predictions, as shown in Table 22, are basically the same as in the previous artificial language learning experiment. However, the participants might be more uncertain about what they are learning so that I generally expected a weaker performance and fewer generalizations than in the previous artificial language learning experiment on retroflexion. As the predictions based on language-general and language-specific perception do not differ for learning, the generalization behavior will be of special interest here. In Table 22 in the second lines the predictions for the case that [ʈ] is perceived as [ɖ] are shown (similar to Table 18 for the previous experiment). Empty cells are those in which the consonant in training would be the one most generalized to. In this case the effect of phonetics and the effect of exposure during the training phase of the experiment would not be separable.

Table 22: Predictions for the artificial language learning experiment testing retroflexion without alternation (2).

	Learning	**Generalizations of**		
		[t] -learners	**[s] -learners**	**[l] -learners**
Language-general perception	[s] > [l], [t] [s] > [l] > [t]	[s] > [l] –		[s] > [t] [s] > [t]
Language-specific perception	[s] > [l], [t] [s] > [l] > [t]	[l] > [s] –	– –	[t] > [s] [t] > [s]

5.2 Experiment

I will now describe how the experiment was designed by focusing on the stimuli, the procedure and the participants. After that I will present the analysis of the experimental results.

5.2.1 Method

Stimuli The stimuli were the same as in the previous artificial language learning experiment on retroflexion – except that in the present experiment all singular forms were excluded from training and test. The artificial language consisted of plural and diminutive forms. The stimuli were constructed from a subset of the German (Wiese, 1996) and Norwegian phoneme inventories (Kristoffersen, 2000). All items were in agreement with the phonotactics of German (Wiese, 1996) – except for the retroflex consonants and [ɾ]. The structure of the stimuli and examples of each grammatical form are illustrated in Table 23. The plurals were pseudowords of the form $C_1V_1C_2$[ɾa] with [ʂ], [ɭ] or [ʈ] as C_2 and the diminutives were pseudowords of the form $C_1V_1C_2$[na] with [s], [l] or [t] as C_2. The consonants [p], [b], [d], [k], [g], [f] and [v] were used in C_1-position. In V_1-position I used the vowels [ɛ], [ɪ], [ɔ] and [ʊ].

Table 23: Structure of the stimuli for the artificial language learning experiment testing retroflexion without alternation.

Form	**C_1**	**V_1**	**C_2**	**Suffix**	**Example**
plural	[p b d k g f v]	[ɛ ɪ ɔ ʊ]	[ʂ ɭ ʈ]	[ɾa]	[bɛɭɾa]
diminutive	[p b d k g f v]	[ɛ ɪ ɔ ʊ]	[s l t]	[na]	[bɛlna]

For the training phase I used 32 items for each group of participants (16 plurals and 16 diminutives). For the subsequent test phase I used 24 stimulus pairs, each consisting of a plural form which conformed to the retroflexion rule (e.g. correct plural [bɛɭɾa]) and one which did not (e.g. incorrect plural [bɛlɾa]).[11] There was an equal number of eight pairs for each of the three consonants. Half of these pairs were part of the training items (n = 4) and half of them were not (n = 4). The complete set of all stimuli used in the artificial language learning experiment on retroflexion is listed in the appendix (see Table 32, 33, 35).

The stimuli were recorded by a male monolingual native speaker of Norwegian. Recording took place at the University of Tromsø. The sampling rate was 44.1 kHz. Stimuli were recorded in the target sentence *Jeg spiste X til middag.* 'I ate X for dinner.' to focus on the stimulus item and to ensure a uniform language environment that allowed the speaker to naturally produce retroflex consonants. After recording all stimuli were sliced out of their context at the nearest zero-crossing point using Praat (Boersma & Weenink, 2019) to avoid irritating noise at the beginning or at the end of the syllables. The intensity of all stimuli was adjusted to 70 dB using Praat (Boersma & Weenink, 2019).

Procedure The procedure was the same as in the artificial language learning experiment testing vowel nasalization (see chapter 2.4) and the first artificial language learning experiment testing retroflexion (see chapter 4.4) – except that the present experiment was run online. For additional information see the description of the procedure in chapter 2.4.1.

The experiment was divided into a perceptual training phase and a perceptual forced-choice test. It was run on the online experiment platform Experigen (Becker & Levine, 2010). Participants were asked to wear headphones and to run the experiment in a quiet place to avoid disturbing noise. The experiment lasted about 5-10 minutes.

Participants were trained on a subset of the stimuli but tested on all stimuli – including stimuli known from training and novel

[11] I decided to focus on plural forms only in the test phase as plural was the crucial condition and as a shorter test phase should avoid that participants stopped the online experiment before it was completed.

stimuli they had not yet heard. There were three training groups: During training each group heard the retroflexion of only one consonant – either [s] or [l] or [t]. For example, a member of the [s]-group never received any training for [l] or [t]. The test was identical for all groups and included all consonants.

In a short introductory phase the participants were familiarized with the set-up of the experiment by listening to German animal names in the plural and in the diminutive. The auditory stimuli were supported visually with images (Snodgrass & Vanderwart, 1980) to illustrate their meaning.

The introductory phase was followed by the training phase and the test phase. In the training phase the participants listened to two repetitions of 32 stimuli (16 plurals and 16 diminutives) in randomized order. An auditory stimulus was played while an image of a fantasy animal (van de Vijver & Baer-Henney, 2014) was shown (compare Figure 12). A plural form was accompanied by an image of two fantasy animals and a diminutive form was accompanied by an image of a small fantasy animal. During training participants received positive input only, which means that they never listened to an incorrect plural, e.g. [bɛlɾa].

The test after training was a perceptual forced-choice task and it was the same for all groups. There were 24 stimulus pairs consisting of a correct and an incorrect plural form differing only in the retroflexion of C_2. There were eight pairs with [s], [l] and [t]. A trial consisted of the presentation of the first form, followed by the presentation of the second form. There was an equal number of trials in which the first and the second form was the correct one. The auditory stimuli were supported visually by an image of two fantasy animals (compare Figure 13). The participants had to decide which of the two forms was the correct one by clicking on a button on the screen. The participants had the choice between buttons labeled with *1* for the first form and *2* for the second form. After their decision the next stimulus pair was presented. All stimulus pairs were presented in random order.

Participants 80 native speakers of German completed the experiment (53 identified as female, 23 identified as male, 4 identified as gender diverse, mean age: 30.5, range: 19-64). They were randomly assigned to one of the three experimental groups. 25 participants were trained with retroflexion of [l], 26 participants were trained with retroflexion of [s], and 29 participants were trained with retroflexion of [t]. All groups were tested with all consonants [l], [s] and [t]. No one reported knowledge of a language which uses retroflex consonants distinctively (except American English). All participants had normal or corrected-to-normal vision and no hearing problems. All of them participated voluntarily. A prize was raffled off for the participants. Each participant agreed to let me use the collected data and filled out a form giving information about the participants' age, sex, native language and foreign language skills.

5.2.2 Results

I analyzed the proportion of correct answers from 1920 data points (80 participants $\times$ 24 trials) with R (R Core Team, 2015) by means of a generalized linear mixed-effects analysis with the corresponding R package lme4 (Bates et al., 2015).

I was interested in the learning and generalization behavior across experimental groups. An overview of the results can be found in Figure 27. This Figure was created using R (R Core Team, 2015) and the R packages Rmisc (Hope, 2013) and ggplot2 (Wickham, 2009). First of all, I compared the results of the trained consonant in each group. This is what I call *learning*. In the model which fitted the data best (as assessed by backward stepwise elimination (Baayen, 2008)) ACCURACY (correct and incorrect) served as dependent variable and TRAINED CONSONANT ([l], [s] and [t]) was the independent variable. PARTICIPANT and ITEM were included as random intercepts (R syntax: glmer(ACCURACY $\sim$ TRAINED CONSONANT + (1|PARTICIPANT) + (1|ITEM), family = "binomial")).

In the trained condition [s]-learners gave with 84% significantly more correct responses to [s]-items than [l]-learners with

65% to [l]-items (Est. = −1.4085, SE = 0.6724, z = −2.095, p < 0.05) and [t]-learners with 41% to [t]-items (Est. = −2.7529, SE = 0.6709, z = −4.103, p < 0.001). [l]-Learners gave significantly more correct responses to [l]-items than [t]-learners to [t]-items (Est = −1.3442, SE = 0.6416, z = −2.095, p < 0.05). [s]-Learners' responses differed significantly from chance level (Est. = 2.2869, SE = 0.4982, z = 4.590, p < 0.001), those from [l]-learners differed marginally significantly from chance level (Est. = 0.8784, SE = 0.4613, z = 1.904, p = 0.057) and those from [t]-learners did not differ from chance level (Est. = −0.4660, SE = 0.4448, z = −1.048, p = 0.295).

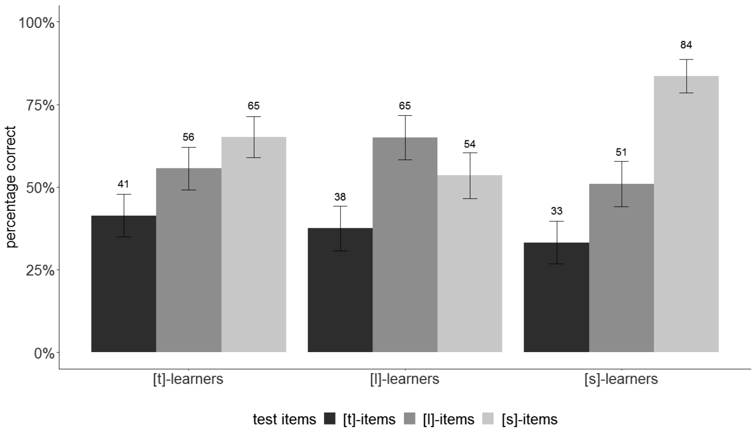

Figure 27: Percentages of correct responses ± 1.96 SE across all training groups and all consonant manners at test (2).

I also investigated how participants judged stimuli with untrained consonants. This is what I call *generalizations*. Therefore I performed three generalized linear mixed-effects model – one for each learning group. In the models which fitted the data best (as assessed by backward stepwise elimination (Baayen, 2008)), ACCURACY (correct and incorrect) served as dependent variable and CONSONANT AT TEST ([l], [s] and [t]) as independent variable. PARTICIPANT and ITEM were included as random intercepts (R syntax: glmer(ACCURACY ~ CONSONANT AT TEST + (1|PARTICIPANT) + (1|ITEM), family = "binomial")).

[t]-Learners responded correctly to [t]-items in 41%, to [s]-items in 65% and to [l]-items in 56%. The percentages of correct responses to [s]-items were significantly above chance level (Est. = 0.7202, SE = 0.3070, z = 2.346, $p < 0.05$), whereas those to [l]-items did not differ significantly from chance level (Est. = 0.2598, SE = 0.3034, z = 0.857, p = 0.392). The percentage of correct responses to [t]-items did not differ significantly from that to [l]-items (Est. = 0.6665, SE = 0.4116, z = 1.619, p = 0.105), but there were significantly less correct responses to [t]-items than to [s]-items (Est. = 1.1268, SE = 0.4145, z = 2.718, $p < 0.01$). The percentages of correct responses to [l]-items and to [s]-items did not differ significantly from each other, either (Est. = 0.4603, SE = 0.4133, z = 1.114, p = 0.265).

[l]-Learners responded correctly to [l]-items in 65%, to [s]-items in 54% and to [t]-items in 38%. The percentages of correct responses to [t]-items was marginally significant below chance level (Est. = −0.5550, SE = 0.2995, z = −1.853, p = 0.064), those to [s]-items did not differ from chance level (Est. = 0.1613, SE = 0.2994, z = 0.539, p = 0.590). There was no significant difference between the percentages of correct responses to [l]-items and to [s]-items (Est. = −0.5807, SE = 0.4124, z = −1.408, p = 0.159). However, there were significantly more correct responses to [l]-items (Est. = −1.2970, SE = 0.4135, z = −3.137, $p < 0.01$) and marginally significantly more correct responses to [s]-items (Est. = −0.7163, SE = 0.4075, z = −1.758, p = 0.079) than to [t]-items.

[s]-Learners responded correctly to [s]-items in 84%, to [l]-items in 51% and to [t]-items in 33%. The percentages of correct responses to [l]-items did not differ significantly from chance level (Est. = 0.0486, SE = 0.3393, z = 0.143, p = 0.886), whereas those to [t]-items were significantly below

chance level (Est. = -0.7837, SE = 0.3403, z = -2.303, $p < 0.05$). There were significantly more correct responses to [s]-items than to [t]-items (Est. = -2.7150, SE = 0.4979, z = -5.453, $p < 0.001$) and to [l]-items (Est. = -1.8827, SE = 0.4943, z = -3.809, $p < 0.001$). There were marginally significantly more correct responses to [l]-items than to [t]-items (Est. = -0.8323, SE = 0.4678, z = -1.779, p = 0.075).

5.3 Discussion

This artificial language learning experiment showed that patterns which do not give evidence for a morphophonological alternation lead to more uncertainty than patterns which give evidence for a morphophonological alternation. Compared to the previous artificial language learning experiment (see chapter 4.4), in which support for an alternation was offered during training by the presentation of singular forms, in the present experiment, in which no support for an alternation was offered during training by omitting singular forms, there were fewer generalizations and more responses below chance level. Nevertheless, two groups of participants learned the retroflexion pattern even without singulars.

As this artificial language learning experiment aimed at shedding light on two different research questions, I will discuss both of them separately. I will start with the discussion of the learning strategies and will then move on to the discussion of the interaction between language-specific and language-general perceptual cues.

Exposing participants not to the base form singular enabled me to investigate whether learners need a relation between forms of a paradigm to accurately generalize their trained pattern as proposed by the relation account or whether they are able to transfer the feature [retroflex] from trained plurals to untrained plurals without knowing about an alternation as proposed by the transfer account. The results in the trained conditions indicate that [s]-learners and [l]-learners were able to learn the pattern they were exposed to since they performed above chance level in their trained conditions. [t]-Learners, however, were not able to learn their pattern – they performed at chance level in their trained condition. As [t]-learners performed at chance level in the previous artificial language learning experiment on retroflexion as well, the reason for the performance at chance might be the same as well: [t]-Learners might have been biased by structural complexity as they perceived [ʈ] as [ɖ] and were thus confronted with a structurally more complex pattern than [s]-learners and [l]-learners. The learning results are hence not only in line with the transfer account but also with the relation account. Table 24 provides a comparison of the predictions and the experimental results. Green check marks represent confirmed predictions and red x marks represent unconfirmed predictions.

Table 24: Predictions for the artificial language learning experiment testing retroflexion without alternation and their evaluation (1).

	Learning	Generalizations of		
		[t] -learners	[s] -learners	[l] -learners
Transfer account	✓✓	✓X	✓X	✓X
Relation account	✓✓	– ✓	– ✓	– ✓

The results in the untrained conditions indicate that no group generalized their trained pattern to untrained items. None of the three learning groups performed above chance level in their untrained conditions. Both, [s]-learners and [l]-learners, performed at chance level or below chance level with untrained items. The only exception are the [t]-learners who performed with [s]-items above chance level. However, as [t]-learners did not even learn their trained rule, this good performance with [s]-items cannot be explained with a transfer of the feature [retroflex] to [s]-items. For this transfer, [t]-learners had to learn their trained rule and they had to perform above chance level with [l]-items as well. Consequently, the results provide no evidence for the transfer account but only for the relation account (see Table 24). Participants seem to need a relation between forms to figure out how the phonological pattern is formed in order to be able to generalize it to novel items. Without a relation between forms a generalization is not possible.

Note, however, that the generalization results might also be explained with different reasons. As in the previous learning experiment the generalization results may be due to structural simplicity for [t]-learners, due to a wrong mapping on an L1-category for [s]-learners and due to a more difficult task than in the previous experiment for [l]-learners (for details see chapter 4.4.3 and below).

The results of this artificial language learning experiment do not only demonstrate how important the relation between forms in a paradigm is for a successful acquisition of a phonological pattern but they also reveal information about the status of phonemes and allophones during learning. Based on the transfer account participants interpret the different consonants in plurals and diminutives as two different phonemes, e.g. [l] and [ɭ]. However, based on the relation account participants interpret the different consonants in (singulars,) plurals and diminutives as allophones of one phoneme, e.g. [ɭ] is an allophone of [l]. The successful learning of the different allophones and the generalization of these patterns in the previous learning experiments compared to the lack of generalizations in the present learning experiment demonstrate that allophones are more important units than phonemes during the process of phonological learning. This result is not surprising considering the work of Mitterer, Sharenborg & McQueen (2013) on perceptual learning (Eisner & McQueen, 2005; Kraljic & Samuel, 2005, 2006; Norris, McQueen & Cutler, 2003).

Mitterer et al. (2013) exposed Dutch native speakers in a lexical decision task to words with an ambiguous final consonant between [ɹ] and dark [ɫ]. In Dutch the consonant [ɹ] is, together with [ʀ] and [r], an allophone of /r/, whereas [ɫ] is, together with [l], an allophone of /l/. The ambiguous consonant replaced either a final /r/ or a final /l/ in existing Dutch words. In a subsequent perceptual categorization task participants were exposed to several tokens of different continua between the allophones of /r/ and /l/: [ɹ]-[ɫ],[r]-[ɫ] or [r]-[l]. The results show that participants who listened to /r/-final words in the lexical decision task categorized the ambiguous consonant more often as /r/ than participants who listened to /l/-final words in the lexical decision task. However, this is only true for participants tested on the [ɹ]-[ɫ]-continuum. Participant tested on a continuum which did not include both allophones presented in the lexical decision task, did not show an effect of exposure. This means that participants do not necessarily abstract to phonemic units during lexical processing but to allophonic units. Experimental studies of Mitterer & Reinisch (2017), Mitterer, Reinisch & McQueen (2018) and Reinisch, Wozny, Mitterer & Holt (2014) yielded similar results. It is important to note though, that in these experiments participants made use of lexical knowledge – the lexical representation of the words – which altered pre-lexical information – the phonological representation of the words (Norris et al., 2003). In the artificial language learning experiments described in this dissertation, however, participants cannot made use of lexical knowledge as all words are part of an artificial language the participants have never heard before, and they are even not asked to do so as they do not have to learn the meaning of the presented words but only the phonological properties of the morphological forms. Nevertheless, the results of the present experiment shows together with those of Mitterer & Reinisch (2017), Mitterer et al. (2013, 2018) and Reinisch et al. (2014) that participants rely on allophones, not on phonemes when learning the pre-lexical phonological representation. Consequently, an allophone seems to be a more important pre-lexical unit than a phoneme in speech perception.

In addition to the investigation of the learning strategy, the present artificial language learning experiment aimed at further investigating the interaction of language-specific (= L1) and language-general perceptual cues (= phonetics) dependent on how the training was designed. Therefore I will continue with a comparison of the results of the present experiment with the results of the previous artificial language learning experiment on retroflexion (see chapter 2.4).

The results in the trained conditions of the present experiment are the same as in the previous artificial language learning experiment (see chapter 4.4.2). [s]-Learners were most successful in their trained condition, followed by [l]-learners. [l]-Learners, however, were only marginally above chance level, which shows that the task was more difficult for them than in the previous experiment in which they performed significantly above chance level. Similar to the results of the previous experiment participants in the [t]-group did not learn their pattern. These learning results are in line with a language-general perception account and with a language-specific perception account. Table 25 provides a comparison of the predictions and the experimental results. Green check marks represent confirmed predictions and red x marks represent unconfirmed predictions. Check marks in brackets are based on marginally significant changes or differences to chance level and are no (correct) generalization of a trained pattern. In the second lines the predictions for the case that [ʈ] is perceived as [ɖ] are shown. Empty cells are those in which the consonant in training would be the one most generalized to. In this case the effect of phonetics and the effect of exposure during the training phase of the experiment would not be separable.

Concerning generalizations, the results differ across groups. [t]-Learners did not learn their rule and were thus not able to generalize it to other consonants – similar to the results of the previous experiment (see chapter 4.4.2). [s]-Learners were most successful in their trained condition. Similar to the results of the previous artificial language learning experiment (see chapter 4.4.2), [s]-learners did not generalize their learned rule to other consonants. The reason for this may again be the mapping of the retroflex fricative onto the postalveolar fricative. This kind of alternation could not be perceived in test pairs with [l] and [t] by the participants. [l]-Learners were most successful in their trained condition as well but – in contrast to the results in the previous artificial language learning experiment, in which they generalized their rule to [t]-items (see chapter 4.4.2) – they did not generalize their learned rule to other consonants. The reason for this might be that [l]-learners were more uncertain about the pattern they learned – due to the missing singular forms – than in the previous artificial language learning experiment in which support for the alternation was offered. The results are thus in line with those from the previous artificial language learning experiment on retroflexion and suggest that participants learned the same pattern in both experiments, however, to a weaker degree in the present experiment.

Table 25: Predictions for the artificial language learning experiment testing retroflexion without alternation and their evaluation (2).

	Learning	**Generalizations of**		
		[t] -learners	**[s] -learners**	**[l] -learners**
Language-general perception	[s] > [l], [t] ✗	[s] > [l] (✓)		[s] > [t] (✓)
	[s] > [l] > [t] ✓	– ✓		[s] > [t] (✓)
Language-specific perception	[s] > [l], [t] ✗	[l] > [s] ✗	– ✓	[t] > [s] ✗
	[s] > [l] > [t] ✓	– ✓	– ✓	[t] > [s] ✗

Comparing the results with the predictions concerning learning I can say that the predictions based on language-general perception and the predictions based on language-specific perception are confirmed – assuming that retroflex [ʈ] was perceived as retroflex [ɖ]. In this case, [t]-learners may again – as in the previous artificial language learning experiment – suffer from a complexity bias as they perceived a two-feature change whereas [l]- and [s]-learners perceived a one-feature change (Moreton, 2008, 2012).

Comparing the results with the predictions concerning generalizations, I can say that only the predictions based on language-general perception are – more or less – confirmed for all learning groups, whereas the predictions based on language-specific perception are only confirmed for [s]- and [t]-learners but not for [l]-learners. [l]-Learners performed at chance level with [s]-items and below chance level with [t]-items. Thus, they were more successful with [s]-items than with [t]-items – although they generalized the correct pattern in neither case. Thus, these results show more evidence for language-general perception than for language-specific perception. [s]-Learners did not generalize at all, which is in line with language-specific perception. There are no predictions based on language-general perception for [s]-learners. [t]-Learners did not learn their rule and consequently did not generalize to other items. This is in line with language-general perception and with language-specific perception.

Comparing the confirmed predictions with those from the previous artificial language learning experiment on retroflexion (see Table 22), I can say that the more difficult task due to the lack of singular forms and thus due to the lack of a morphophonological alternation led to more evidence for language-general perception based on perceptual ease and thus for a phonetic bias. Hence, the more difficult it is to figure out the morphophonological pattern, the more participants rely on ease of perception and the fewer participants rely on L1. However, there is also evidence for language-specific perception: As [l]-learners did not generalize – probably due to the more difficult task – and as [t]-learners did not generalize neither, just the generalization results – without looking at percentages below chance – are in line with language-specific perception (those without check marks in brackets in Table 25). As neither of the groups generalized their trained pattern, this is, however, not very convincing.

The results of the present experiment are in line with the assumptions based on the results of Baer-Henney et al. (2015a). In Baer-Henney et al. (2015a)'s experiment participants only relied on phonetics while making generalizations when training was shortest and contained the most exceptions compared to other conditions. Thus, phonetics may be the reliable source in an uncertain learning environment. The more input participants get and the more time they have to figure out what they are exposed to, the more they use other sources to explain the patterns being exposed to. One of these other sources seems to be L1 as the results of Baer-Henney et al. (2015a), Greenwood (2016) and the previous artificial language learning experiment (see chapter 4.4) suggest. In Greenwood (2016)'s experiments the results were affected by the distribution of alternating sounds in the L1-lexicon when the items were presented in careful – reliable – speech during training. However, when the items were presented in casual – less reliable – speech during training, the distribution of alternating sounds in the L1-lexicon did not affect the results at test.

This assumption is in line with van de Vijver & Baer-Henney (2014). In this study van de Vijver & Baer-Henney (2014) tested the plural formation of German five-year-old and seven-year-old children as well as of German adults with a wug test (Berko Gleason, 1958). German plurals can alternate in voicing, which means that singulars ending in a voiceless obstruent, e.g. [ta:k] *Tag* 'day', have plurals with a voiced obstruent, e.g. [taːgə] *Tage* 'days'. Note that the voiceless obstruent in the singulars is due to final devoicing (Wiese, 1996). Moreover, there are German singulars with a back vowel, e.g. [kuː] *Kuh* 'cow', which is a front vowel in the plural, e.g. [kyːə] *Kühe* 'cows'. Whereas the voicing alternation is phonetically motivated, the vowel alternation is not phonetically motivated, but both alternations occur equally often in German. The results revealed that adults produced voicing alternations and vowel alternations in their plurals of pseudowords equally often, which corresponds to the equal distribution of both alternations in the German lexicon. Adults thus relied on their L1-knowledge when forming generalizations. The children, however, produced more plurals with voicing alternation (phonetically motivated) than with vowel alternation (not phonetically motivated), whereas five-year-old children produced more voicing alternations than seven-year-old children. van de Vijver & Baer-Henney (2014) concluded that the smaller the lexicon, and thus the more uncertain children are about the lexical distribution of those alternations, the less they rely on the lexicon and the more they rely on phonetic grounding when forming generalizations.

Putting these results together, it seems as if phonetics is the first reliable source when making generalizations. However, when other L1-specific sources are available, such as frequency distributions in the L1-lexicon, phonetics is downgraded and L1 forms the most reliable source when generalizing. This is in line with results in artificial language learning experiments with adults, e.g. Baer-Henney & van de Vijver (2012), Finley (2012) and Wilson (2006), and with the artificial language learning experiment on vowel nasalization in chapter 2.4, in which no L1-support was offered. If there is no support from L1, even adults are biased by phonetics.

Independent of the generalizations of the trained pattern, the results of the present learning experiment show nevertheless an influence of L1-phonotactics, as other observations in the results show which could not be found in the previous artificial language learning experiment. Both, [l]-learners and [s]-learners responded significantly below chance level to [t]-items. This means that they were certain about this incorrect response. They preferred alveolar [t] in plurals. This may be explained by how participants perceived the retroflex [ʈ] (see Figure 15). Participants were uncertain as to what exactly this retroflex [ʈ] was. Some identified it as [t] (8%), some as [ʈ] (27%), but more than half of the participants identified it as [d] (13%) or as [ɖ] (52%). As the German language does not allow voiced obstruents at the end of a syllable (due to the process of final devoicing), a [d] or [ɖ] in the plural, e.g. [dɪɖ-ɾa] would violate the phonotactics of German (Wiese, 1996). Thus, the preference for [t] over retroflex [ʈ] (perceived as [d] or [ɖ]) represents a preference for the phonetic and phonological details of L1.

There is one further observation in the results which could not be found in the previous artificial language learning experiment either: [t]-Learners had a preference for retroflex [ʂ] over alveolar [s]. This preference could not be due to training, as [t]-learners did not learn their pattern at all. Hence, no generalizations were possible. Contrary to the preference for alveolar [t] by [l]-learners and [s]-learners, this preference, however, seems not to be due to the German language. If this was the case and assuming that retroflex [ʂ] was perceived as postalveolar [ʃ], I would expect that [ʃ] occurs more often than [s] at the end of a syllable. I checked the BAStat corpus (Schiel, 2010) (online available: https://www.bas.uni-muenchen.de/Bas/BasPHONSTATeng.html), a database with statistics of conversational German based on speech corpora containing spontaneous, conversational speech only, for the occurrences of [s] and [ʃ] in word-final position. Unfortunately, the database does not give information about the occurrences in syllable-initial or syllable-final position. I based my analysis on all available sub-corpora which are *Verbmobil 1 + 2, Regional Variants of German 1* and *SmartKom*. It turned out that [s] is used more often than [ʃ] word-finally in German. Consequently, this cannot explain the preference for retroflex [ʂ] (= postalveolar [ʃ]) over alveolar [s] by [t]-learners. I therefore suggest that the preference for retroflex [ʂ] over alveolar [s] is accidental, especially as this preference is not present in the previous artificial language learning experiment (see Figure 26).

5.4 Conclusion

The present artificial language learning experiment aimed at testing whether participants are able to learn and generalize a phonological pattern without being exposed to a morphophonological alternation and whether differences in training lead to different performances at test. Whereas some phonological theories assume that learners can readily transfer a certain feature, e.g. [retroflex], from one item to a novel one (Berent, 2013; Berent et al., 2012), others assume that learners rely on the relation between phonological forms, e.g. allophonic relations, during speech perception and phonological learning. Besides, previous studies suggested that the amount of exposure and the length of training affect the performance, whereas longer training phases and more amount of exposure generally lead to better performances than shorter training phases and less amount of exposure (Baer-Henney et al., 2015a; Baer-Henney & van de Vijver, 2012). More interestingly, former investigations also showed that differences in training can also lead to different performances depending on the influencing factors L1 and phonetics (Baer-Henney et al., 2015a; Greenwood, 2016). The shorter the training, the less the amount of exposure to an alternation, the more exceptions in training and the less clear the pronunciation of training items – thus, the less reliable the training set – the less is learning affected by L1 and the more is learning affected by phonetics.

I conducted an artificial language learning experiment on the basis of the experiment described in chapter 4.4 but with a modified training set which did not contain singular forms. Participants were exposed to plurals and diminutives only, so that there was no evidence for an alternation, hence a relation, in the items. This was avoided in the previous artificial language learning experiment by the additional exposure to singulars.

The results of the present experiment show that the lack of presentation of the base form singular in the training led to fewer generalizations and more responses below chance level than the training set with singular forms in the previous artificial language learning experiment. [s]-Learners performed best in their trained condition, followed by [l]-learners. [t]-Learners did not learn their pattern at all. These learning results can again be explained not only by the language-general perception account but also by the language-specific perception account. Hence, the learning results cannot be used to differentiate between both accounts. The generalization results, however, can be used to differentiate between them: They can be better interpreted in the light of the language-general perception account: Whereas [s]-learners did not generalize at all, [l]-learners performed better with [s]-items than with [t]-items and [t]-learners performed better with [s]-items than with [l]-items. Note, however, that [l]-learners did not generalize above chance level and that [t]-learners did not learn their trained pattern. The performance concerning the generalization can thus be interpreted in that way that a training set without evidence for an alternation leads to a greater influence of phonetics and to a smaller influence of L1, which is in line with the expectations based on former studies (Baer-Henney et al., 2015a; Greenwood, 2016). The generalization results further demonstrate that learners fail to generalize a phonological pattern by means of pure transfer. Learners need information about the phonological relation of different forms of a paradigm to accurately predict the morphophonological form of novel items. This speaks against the transfer account based on Berent (2013) and Berent et al. (2012) but is in line with the findings of Mitterer & Reinisch (2017), Mitterer et al. (2013, 2018) and Reinisch et al. (2014) that allophones are more important units than phonemes during speech perception.

6 General discussion

The general discussion is structured as follows: At first I will summarize the findings of my experimental studies, then I will describe for each factor under investigation how it affects learning and generalizing a morphophonological pattern. In the end I will link my findings and propose how phonetics and especially perception shapes phonological learning.

6.1 Summary of the experimental results

In this dissertation I studied the role of phonetics in learning phonological alternations. I focused on how different phonetic biases interact and how different tests and different trainings influence the performance by studying the learning behavior and the generalization behavior of native German participants. The results of former studies investigating a phonetic bias were sometimes confounded by a simplicity bias, e.g. Cristià & Seidl (2008) and Peperkamp et al. (2006), or by phonological characteristics of the participants' native language, e.g. Baer-Henney (2015), chapter 4. In order to avoid such confounds, I decided to expose participants to patterns which differ in phonetic grounding but are structurally akin, and I decided to include the perceptual similarity to L1-phoneme categories as factor. To this end, participants were trained on a vowel nasalization pattern differing in vowel height and to a retroflexion pattern differing in manner of articulation. German native speakers are neither familiar with phonemic vowel nasalization nor with retroflexion. The artificial language learning experiments were framed with short typological surveys investigating the distribution of the trained sound patterns in the world's languages as well as with a production experiment, an acoustic analysis and perception experiments to establish which of the trained sound patterns are articulatory easy, which are perceptually easy, which are perceptually similar to each other and which are perceptually similar to native sound categories for German native speakers.

The first experiments focused on vowel nasalization. According to the literature (Bell-Berti, 1993; Ohala, 1975; Schwartz, 1968), low vowels are articulatory easier to nasalize than high vowels, whereas high oral and nasal vowels are perceptually easier to distinguish than low oral and nasal vowels. Chapter 2 starts with a short survey on the typological distribution of high, mid and low oral and nasal vowels. It turned out that high, mid and low nasal vowels are equally distributed across the world's languages, which means that the two predispositions based on ease of perception and ease of articulation are not the – or not the only – reason for the equal distribution of all nasal vowel heights across languages. If this was the case, I would have expected a lower number of mid nasal vowels compared to high and low nasal vowels in the languages of the world. The following short acoustic measurement confirmed the acoustic proximity of (lower) mid [ɛ] to low [a] and its acoustic distance to high [i] – independent of nasality, which is important when investigating the phonetic similarity as factor affecting generalizations in phonological learning.

The production experiment in chapter 2.2 and the subsequent acoustic analysis investigated which vowels German native speakers nasalize in nasal contexts. It turned out that German native speakers nasalize non-high vowels more heavily than high vowels. This means that German native speakers benefit from ease of articulation: The lower the vowel, the easier it was to nasalize. The perception experiment in chapter 2.3 investigated the confusability of oral and nasal vowels by means of an identification task. The results confirmed the perceptual similarity of high oral and nasal vowels and the perceptual distance of low and mid oral and nasal vowels. They further confirmed the acoustic and thus perceptual similarity of low [a] and lower mid [ɛ] as well as the perceptual distance of lower mid [ɛ] and high [i]. This means that German native speakers benefit from ease of perception: The higher the nasal vowel, the easier it was to perceive.

In the artificial language learning experiment (see chapter 2.4) three groups of participants were trained on the nasalization of either high or mid or low vowels, and were then tested on the nasalization of all three vowel heights. Doing so, I tested whether learning is more strongly biased by ease of perception or by ease of articulation. The results suggest that learning is biased by ease of perception. Concerning the generalizations to untrained vowel heights, I tested whether learners generalize more to a perceptually easy pattern, to an articulatory easy pattern or to a phonetically similar pattern. I found that learners generalized more to a phonetically and perceptually similar pattern than to a perceptually or articulatory easy pattern.

To sum up, the experiments in chapter 2 provide evidence for a phonetic bias based on perception in learning and for a phonetic bias based on phonetic similarity in generalizations (for a definition of these terms see chapter 1.1). As this phonetic similarity is also found in perception, one could follow that perception shapes morphophonological learning.

The follow-up experiment in chapter 3 was designed to test whether the great influence of perception in the previous artificial language learning experiment (see chapter 2.4) may be due to the perceptual training and the perceptual test. Consequently, this follow-up experiment tested the articulation. It consisted of two reading sessions with French loan words and an auditory exposure to vowel nasalization in between the two reading sessions as well as of an acoustic analysis of the recordings. Participants were exposed to the nasalization of either high or mid or low vowels. I tested whether learning is more strongly affected by ease of perception, ease of articulation or perceptual frequency. The results were variable, but suggest that learning is biased by ease of perception and perceptual frequency. Concerning the generalizations to untrained heights, I tested whether speakers generalize more to a perceptually easy pattern, to an articulatory easy pattern or to a phonetically similar pattern. I found that speakers, if at all, generalized to a phonetically similar pattern.

The experiments on vowel nasalization (chapters 2-3) thus showed that learning is biased more by perception than by articulation. This is in line with Glewwe (2019) who claimed that, if a phonetic bias exists at all, it will be based on perception. The results further show that generalizations are affected by phonetic similarity and not by ease of perception or ease of articulation. This effect is independent of the task, whereas it is stronger in a perceptual test than in a productive test. This is in line with the perception before production hypothesis (Flege, 1991), which states that learners can perceive contrasts before they can articulate them. In general, I found an effect of ease of perception in synchronic phonology. However, an effect of ease of perception on typological markedness could not be found.

The second experiments focused on retroflexion and the influence of different aspects of perception in phonological learning. These aspects are the perceptual similarity between alternating sounds as well as the perceptual similarity of retroflex consonants to native language sounds. Fricatives unite both aspects: In fricatives the acoustic differences are more salient than in other consonants (Kohler, 1990), meaning a perceptual advantage for fricatives, and a retroflex fricative is perceptually very similar to the German postalveolar fricative. Chapter 4 started with a short survey on the typological distribution of retroflex consonants. It turned out that retroflex plosives are more frequent than retroflex fricatives and retroflex lateral approximants. Retroflex nasals are even less frequent across languages in the world. This means that the predisposition based on ease of perception (and the ones of ease of articulation, too) is not the – or not the only – reason for the asymmetric distribution of retroflex consonants cross-linguistically. If this was the case, I would have expected a higher number of retroflex fricatives compared to other consonant manners in the languages of the world.

The first perception experiment in chapter 4.2 investigated the perceptual confusability of retroflex and alveolar consonants by means of an identification task. The results confirm the perceptual dissimilarity of retroflex and alveolar fricatives and suggest a perceptual advantage for retroflex consonants in VC compared to CV. This means that German native speakers benefit either from ease of perception – as differences among fricatives are more salient than in other consonants – or from the perceptual similarity to native language phonemes – as a mapping of retroflex fricatives onto postalveolar fricatives would make the fricative contrast a native one. The second perception experiment in chapter 4.3 investigated the discriminability of alveolar, retroflex and postalveolar consonants. The results confirm the perceptual similarity of retroflex and postalveolar fricatives and show that the contrast between retroflex and alveolar fricatives as well as the contrast between retroflex and postalveolar fricatives is better discriminated than the contrasts between retroflex and alveolar plosives or lateral approximants.

In the artificial language learning experiment in chapter 4.4 three groups of participants were trained on the retroflexion of either fricatives or plosives or lateral approximants, and were then tested on the retroflexion of all three consonant manners. Doing so, I tested whether learning and generalizations are more strongly biased by language-general ease of perception or by language-specific perception. The results suggest that learning is biased by perception. Whether this is due to language-general or language-specific factors cannot be stated as both factors predict the same – a learning advantage for fricatives. Concerning the generalizations, I found that the learners' generalizations are based on language-specific factors. I assume that the retroflex fricative was perceived as native postalveolar fricative by German native speakers. Thus, participants exposed to retroflex fricatives did not learn a retroflexion pattern but perceived a native contrast in the stimuli. Consequently, these participants were not able to generalize their learned rule to other stimuli as these other stimuli did not include an alveolar-postalveolar contrast. In addition to that, participants exposed to other consonants were not able to generalize their learned retroflexion rule to fricatives, as the stimuli with fricatives were perceived as stimuli with an alveolar-postalveolar contrast and not with an alveolar-retroflex contrast.

To sum up, the experiments in chapter 4 provide evidence for a perception based bias in learning, which is either based on language-specific or on language-general factors, and for a strong influence of the perceptual similarity to native language phonemes in morphophonological learning.

The follow-up experiment in chapter 5 was designed to test on the one hand whether learners can transfer a phonological pattern to novel items without being exposed to a morphophonological alternation and on the other hand whether the strong influence of perceptual similarity to native language phonemes in the artificial language learning experiment in chapter 4.4 may be due to the reliable training which provided direct evidence for a morphophonological alternation. The second assumption is based on former studies (Baer-Henney et al., 2015a; Greenwood, 2016) which showed that participants relied more on phonetics and less on L1, the less reliable training was. Consequently, this follow-up experiment included a training which did not provide evidence for a morphophonological alternation by omitting singular forms. Participants were exposed to plural forms and diminutive forms which showed the retroflexion of either fricatives or plosives or lateral approximants. I tested whether learning and generalizations are more strongly biased by language-general ease of perception or by language-specific perception and whether generalizations are possible when there is no evidence for a morphophonological alternation. The learning results were similar to the previous artificial language learning experiment in chapter 4.4 and suggest that learning is biased by perception. Whether this is due to language-general or language-specific factors cannot be stated. Concerning the generalizations, I found fewer generalizations and more performances at chance than in the previous artificial language learning experiment, which in general suggest that the learners' performances with untrained items are based on language-general factors when training provides less or no evidence for an alternation. As actually no group generalized their trained pattern to untrained items, the results further demonstrate that learners cannot transfer a phonological patterns to novel items without knowing about the relations between phonological forms.

The experiments on retroflexion (chapters 4-5) thus showed that learners need knowledge about the phonological relation between forms of a paradigm to accurately predict novel items. This is against the assumption of Berent (2013) and Berent et al. (2012). The experiments further demonstrate that the perceptual similarity to native language phonemes influences learning and generalizing a phonological pattern. The more evidence for a morphophonological alternation during training is provided, the more influence L1 has on the performance. The less evidence for a morphophonological alternation during training is provided, the more influence phonetics has on the performance. This is in line with the studies of Baer-Henney et al. (2015a) and Greenwood (2016). Together with the results of the former studies, my results suggest that phonetics is the first reliable source used during phonological learning. It is even used by children and infants who do not have a great knowledge of their L1-lexicon or their L1-phonotactics (van de Vijver & Baer-Henney, 2014). However, as soon as other sources are available, e.g. native language, phonetics is downgraded and L1 is upgraded. Nevertheless, I found an effect of ease of perception in synchronic phonology, at least when evidence for a morphophonological alternation was lacking. However, an effect of ease of perception on typological markedness could not be found.

Connecting the results of the experiments on vowel nasalization and retroflexion with each other, I can say that learners are biased by phonetics in morphophonological learning. Their learning behavior and their generalization behavior is strongly biased by phonetics, especially by perception. This bias is grounded in perceptual ease for learning and in phonetic similarity for generalizations. A role of articulatory ease during phonological learning could not be confirmed. I further found that the perceptual similarity to native language sounds influences how participants learn and generalize. If this is the case, phonetics is downgraded and assumed to be less reliable than the knowledge about the native language. In addition to the results concerning learning biases, the experiments indicate that learners cannot transfer a phonological feature of one item to a novel item without knowing about the relation between these items. They thus highlight the importance of phonological alternations and the corresponding relation between allophones during the perception of speech.

6.2 Phonetic factors affecting phonological learning

I will now describe for each phonetic factor under investigation how it affects the learning and generalization behavior of the participants. Doing so, I will evaluate which of the factors more strongly influence learning and which of the factors more strongly influence generalizations. Further, I will evaluate which phonetic factors form stronger biases than the other factors. For each phonetic factor I will evaluate its role in learning and in generalizations separately.

6.2.1 The role of ease of articulation

I studied the role of a phonetic bias based on ease of articulation (based on the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006)) in the experiments on vowel nasalization. In a production experiment I found that German native speakers coarticulatory nasalized vowels in the context of a nasal consonant (VN or NV) more heavily when these vowels were low or mid than when they were high. High vowels, however, were not coarticulatory nasalized in the context of nasal consonants (see chapter 2.2).

German native speakers thus benefit from ease of articulation. The nasalization of low vowels is articulatory easier than the nasalization of high vowels, as the muscles used to lower the velum – to nasalize a vowel – and the muscles used to lower the tongue body – to produce a low vowel – are anatomically connected (Bell-Berti, 1993; Ohala, 1975).

If this ease of articulation is active during phonological learning, one will expect that it is easier to learn a nasalization pattern in which low vowels are nasalized than a nasalization pattern in which high vowels are nasalized. Further, one will expect that learners, once a nasalization pattern is acquired, will generalize it to the articulatory easy pattern. I tested this by means of an artificial language learning experiment (see chapter 2.4), but I found neither evidence for a learning advantage for low vowel nasalization nor a generalization advantage for low vowel nasalization – except for [ɛ]-learners who generalized their nasalization pattern more to low vowels than to high vowels. However, phonetic similarity makes the same prediction for [ɛ]-learners as ease of articulation. As the predictions based on phonetic similarity are confirmed for all learning groups – not only for [ɛ]-learners – the confirmed prediction based on ease of articulation for [ɛ]-learners may thus be accidental.

As the lack of an effect of ease of articulation in this artificial language learning experiment may be due to the perceptual test, I conducted a follow-up experiment, in which I measured the modification of the amount of nasality in vowels in French loan words before and after exposure to either high or mid or low vowel nasalization (see chapter 3.2). If speakers are biased by ease of articulation, I will expect the greatest increase of the amount of nasality for those participants who were exposed to low vowel nasalization as well as a greater increase of the amount of nasality of low vowels than of non-low vowels by the other learning groups. However, even in a productive test neither a learning advantage nor a generalization advantage for low vowel nasalization did show up. None of the predictions based on ease of articulation were confirmed.

I thus have to conclude that I did not find any evidence for the existence of an articulatory bias during learning. This is in line with what was found by (Glewwe, 2019). In her experiments, a phonetic bias was always weak, and when it showed up, it was always based on perceptual ease.

It may be the case that the use of French loan words may be the reason why I did not find an effect of ease of articulation during learning. Maybe speakers are less able to adjust the pronunciation of a loan word than of a native word, as they are generally uncertain about how to pronounce these loan words. It may be the case that they pronounce the French loan words in the way they have heard it before from other speakers and do not dare to change it – especially not in a linguistic experimental set-up. It may thus be the case that I would have found other effects – possibly in line with a phonetic bias based on ease of articulation – if I had measured the modification of the amount of nasality in vowels in the context of nasal consonants in German words or in pseudowords. As German native speakers are biased by ease of articulation during coarticulatory nasalization of pseudowords (see chapter 2.2), a modification in the amount of nasality might be observable in pseudowords after exposure to vowel nasalization. Nevertheless, as the stimuli in the artificial language learning experiment with French loan words were not counterbalanced and as the results were highly variable it may even be the case that the design of the experiment was not suitable to investigate whether speakers adjust the amount of nasality after exposure to vowel nasalization depending on vowel height. Solving theses issues will be left for future work.

6.2.2 The role of ease of perception

I studied the role of a phonetic bias based on ease of perception (based on the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006)) in the experiments on vowel nasalization and in the experiments on retroflexion. In perception experiments I found on the one hand that oral and nasal high vowels were less often confused with each other than oral and nasal mid or low vowels (see chapter 2.3) and on the other hand I found that alveolar and retroflex fricatives were less often confused with each other than alveolar and retroflex plosives, lateral approximants or nasals by German native speakers (see chapter 4.2).

German native speakers thus benefit from ease of perception. The perception of high nasal vowels is perceptually easier than the perception of low nasal vowels as the acoustic and, thus, the perceptual contrast between oral and nasal high vowels is greater than that of oral and nasal low vowels. This is due to the different degree of acoustic modifications during the nasalization, which means that mid and low oral and nasal vowels are acoustically more similar to each other than high oral and nasal vowels (Schwartz, 1968). In the case of retroflexion the reason lies in the acoustics as well. As fricatives are generally acoustically more salient than other consonant manners, modifications of fricatives are perceived better than modifications of other consonant manners (Kohler, 1990).

If this ease of perception is active during phonological learning, one will expect that it is easier to learn a nasalization pattern in which high vowels are nasalized than a nasalization pattern in which low vowels are nasalized. Further, one will expect that learners, once acquired a nasalization pattern, will generalize it to the perceptually easy pattern. I tested this by means of an artificial language learning experiment (see chapter 2.4) and I found a learning advantage for high vowel nasalization, but no generalization advantage for high vowel nasalization. I further tested for a bias based on ease of perception by means of a productive test with French loan words (see chapter 3). If speakers are biased by ease of perception, I will expect the greatest increase of the amount of nasality for those participants who were exposed to high vowel nasalization as well as a greater increase of the amount of nasality of high vowels than of non-high vowels by the other exposure groups. I found again a learning advantage for high vowel nasalization but no generalization advantage for high vowel nasalization. However, as perceptual frequency (see chapter 6.2.3) predicts the same outcome for learning, it cannot be stated whether this learning advantage in production is due to ease of perception, perceptual frequency or both of them.

Figure 28 illustrates how ease of perception affects learning vowel nasalization. Note that for illustrative purposes it is assumed that ease of perception is the only factor affecting learning. Oral and nasal high vowels are perceptually more distant from each other than oral and nasal mid and low vowels, which is illustrated with the different distances of the circles in the Figure. These circles represent the native sound categories (on top) – the oral vowels – and the learned sound categories (below) – the nasal vowels. The differences in distance between oral and nasal vowels are the reason why the alternation from an oral vowel to a nasal vowel is learned better for high vowels than for mid and low vowels. In this Figure arrows reflect the learning of the unfamiliar, non-native nasal vowels (circles) by means of auditory exposure to nasal vowels (rectangles). Thick arrows mean a better learning than thin arrows. Nasal high vowels are learned best, which is illustrated with the thick circle around the high nasal vowel in comparison to the thin circles around mid and low nasal vowels.

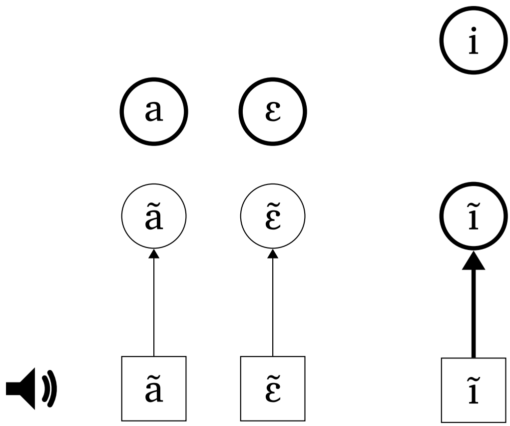

Figure 28: Learning vowel nasalization based on ease of perception.

Concerning retroflexion, one will expect that it is easier to learn a retroflexion pattern in which fricatives are retroflexed than a retroflexion pattern in which plosives or lateral approximants are retroflexed. Further, one will expect that learners, once acquired a retroflexion pattern, will generalize it to the perceptually easy pattern. I tested this by means of an artificial language learning experiment (see chapter 4.4) and I found a learning advantage for retroflex fricatives, but no generalization advantage for retroflex fricatives – except for better performances with fricatives than with plosives or lateral approximants when there was no evidence for an alternation in the artificial language learning experiment in chapter 5. For a discussion of the lack of evidence of an alternation in the training see chapter 6.3.2. However, perceptual similarity to native language phonemes (see chapter 6.2.5) predicts the same outcome for the learning behavior as ease of perception does, so that it is not clear whether the learning advantage for retroflex fricatives is due to ease of perception or due to the perceptual similarity to the German postalveolar fricative or due to a combination of them.

Figure 29 illustrates how learning retroflexion is affected by ease of perception. Note that for illustrative purposes it is assumed that ease of perception is the only factor affecting learning. Alveolar and retroflex fricatives are perceptually more distant from each other than alveolar and retroflex plosives or lateral approximants, which is illustrated with the different distances of the circles in the Figure. These circles represent the native sound categories (on top) – the alveolar consonants – and the learned sound categories (below) – the retroflex consonants. The differences in distance between alveolar and retroflex consonants are the reason why the alternation from an alveolar consonant to a retroflex consonant is learned better for fricatives than for lateral approximants and plosives. In this Figure arrows reflect the learning of the unfamiliar, non-native retroflex consonants (circles) by means of auditory exposure to retroflex consonants (rectangles). Thick arrows mean a better learning than thin arrows. Retroflex fricatives are learned best, which is illustrated with the thick circle around retroflex fricatives in comparison to the thin circles around retroflex lateral approximants and plosives.

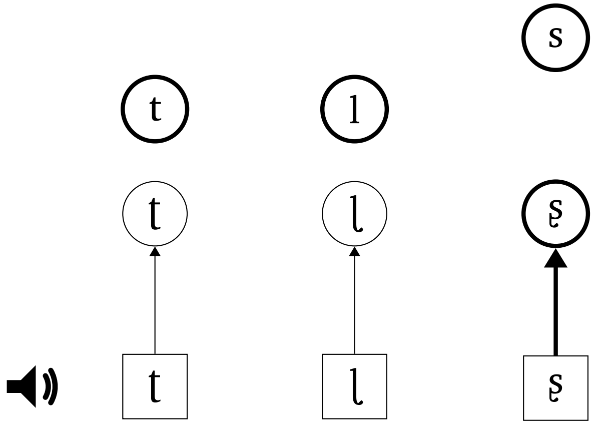

Figure 29: Learning retroflexion based on ease of perception.

Although in the case of vowel nasalization not only ease of perception but also perceptual frequency and in the case of retroflexion not only ease of perception but also the perceptual similarity to native language phonemes can explain the learning behavior of the participants, it is always perception that shapes phonological learning in these experiments. Not only ease of perception but also the perceptual similarity to native language phonemes as well as perceptual frequency refer to how we perceive the pattern we are exposed to during training.

Glewwe (2019) argued for a role of perception in phonological learning as well. In her experiments she found evidence for a phonetic bias only when this bias was based on perceptual ease and not when it was based on articulatory ease. She thus concluded, that if an articulatory based phonetic bias had affected learning in her experiments, its effect would have been too weak so that a stronger bias, in her case a simplicity bias, might have overridden its effect. Glewwe (2019) further suggested breaking the phonetic bias down into subtypes – a perceptually based phonetic bias and an articulatory based phonetic bias. This is exactly what I did in the artificial language learning experiment on vowel nasalization (see chapters 2.4 and 3.2) as well. Based on her results and on a review of the literature on artificial language learning – my results also confirm it – she claimed that only a perceptually based phonetic bias affects phonological learning. According to Glewwe (2019) the majority of studies found a learning advantage for patterns that are easy to perceive (Carpenter, 2006; Finley, 2012; Finley & Badecker, 2012; Greenwood, 2016; Kimper, 2016; White, 2013; Wilson, 2006), whereas studies aimed at finding a learning advantage for patterns which are easy to articulate failed to find differences in learning or had problematic designs (Baer-Henney, 2015; Peperkamp & Dupoux, 2007; Pycha et al., 2003; Saffran & Thiessen, 2003; Seidl & Buckley, 2005; Skoruppa & Peperkamp, 2011).

I think that the review of the literature may be a hint at a stronger effect of a perceptually based bias but I do not consider it a strong argument as it is difficult to design an experiment in a way that it really tests what it is supposed to test. There are many factors that might influence the learning behavior of the participants which cannot always be controlled for. Further it is difficult to decide what makes a pattern easy to perceive or to articulate and finding patterns that differ in only one of these factors can be hard. In addition to that, null results are hardly published, so that there might be many more – unfortunately unpublished – studies which did not find support for a perceptually based phonetic bias than assumed. Nevertheless, I agree – mainly based on my own results – with Glewwe (2019) that a perceptually based phonetic bias is stronger than an articulatory based phonetic bias.

6.2.3 The role of perceptual frequency

I studied the role of perceptual frequency in the experiment on vowel nasalization with French loan words (see chapter 3).

To briefly recapitulate: Perceptual frequency refers to how often we are exposed to a phonetic detail and thus, how many memory traces there are for a given phonetic detail (Goldinger, 1998). In the case of vowel nasalization high nasal vowels are low-frequency primes for German native speakers, whereas mid and low nasal vowels are high-frequency primes for German native speakers. This is because German native speakers are exposed to nasal vowels most commonly in French loan words. In French there are no high nasal vowels but only mid and low nasal vowels (Fagyal et al., 2006), so that German native speakers are said to have no experience with high vowel nasalization but more experience with mid and low vowel nasalization based on French loan words.

If perceptual frequency affects learning vowel nasalization, one will expect better imitations of nasalization after exposure to low-frequency primes (= high vowel nasalization) than after exposure to high-frequency primes (= mid/low vowel nasalization). This is exactly what I indirectly found. Speakers exposed to high vowel nasalization did not adjust their amount of nasality in the trained pattern after exposure but decreased the amount of nasality in the untrained patterns, which indirectly increased the amount of nasality in the trained pattern (high vowels) – meaning that the imitation of nasalization after exposure to low-frequency primes was improved in the trained condition. Speakers exposed to mid and low vowel nasalization neither adjusted their amount of nasality in the trained pattern nor in the untrained patterns after exposure – meaning that the imitation of nasalization after exposure to high-frequency primes was not improved in the trained conditions. Perceptual frequency thus seems to affect phonological learning. However, as ease of perception (see chapter 6.2.2) predicts the same outcome, it cannot be stated whether the learning advantage for high vowel nasalization is due to perceptual frequency, ease of perception or both of them.

Figure 30 illustrates how perceptual frequency might affect learning a vowel nasalization pattern. Note that for illustrative purposes it is assumed that perceptual frequency is the only factor affecting learning. The memory traces already present are represented by the bended arrows connecting the rectangles (exposure sounds) with the circles (existing sound categories). Weak memory traces are represented by dotted arrows and strong memory traces are represented by solid arrows. Due to exposure, new memory traces are established, which do not differ in their strength as the amount of exposure is the same for all exposure groups. These are represented by the straight arrows connecting the rectangles with the circles. The difference between the memory traces, which are already present, and those, which are established due to exposure, is greater for high nasal vowels than for mid and low nasal vowels. Thus, the learning effect is greater for high vowel nasalization than for mid and low vowel nasalization.

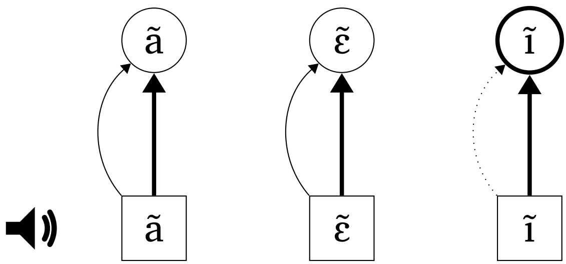

Figure 30: Learning vowel nasalization based on perceptual frequency.

There are no predictions for the generalization behavior based on perceptual frequency as perceptual frequency refers to the phonetic experience before the beginning of the experiment. Participants who are exposed to a pattern with less or weak memory traces benefit most from exposure, whereas participants who are exposed to a pattern with more or stronger memory traces benefit less from exposure. There is, however, no connection between memory traces to high, mid and low vowels assumed, and thus no generalization effect can be predicted based on perceptual frequency.

To my knowledge perceptual frequency has not been investigated as potential factor affecting phonological learning. The lack of predictions for generalizations may be the reason for this. Nonetheless, the role of perceptual frequency in imitations has been studied e.g. by Nielsen (2011) and Zellou et al. (2016, 2017). Although my results are highly variable, my interpretation is in line with their results.

My results further add to the question in how far perception and production are linked to each other. Beddor, Coetzee, Styler, McGowan & Boland (2018) studied the individual perception and production of coarticulatory vowel nasalization by native speakers of American English by means of airflow data, acoustic measurements and a perceptual eye-tracking experiment. In the eye-tracking experiment participants listened to words which either have coarticulatory nasalized vowels as in *bend* or which have no coarticulatory nasalized vowels as in *bed* while seeing pairs of images showing a bed and someone bending a wire or a sheet. When exposed to e.g. *bend*, they found that those participants who produced heavier coarticulatory vowel nasalization looked earlier to the image showing someone bending a wire or a sheet rather than to the image showing a bed than participants who produced weaker coarticulatory vowel nasalization. Beddor et al. (2018) conclude that "an individual's perception is made public through their productions" (p. 931). The direction, however, is according to them not clear, and my results suggest that not only our production can alter how we perceive vowel nasalization but also that our perception can alter how heavy we produce vowel nasalization.

6.2.4 The role of phonetic similarity

I studied the role of phonetic similarity (based on Kapatsinski (2013a)) in the experiments on vowel nasalization. In a short acoustic analysis I found that lower mid [ɛ] is acoustically more similar to low [a] than to high [i] – independent of whether these vowels are oral vowels or nasal vowels (see chapter 2.1).

This acoustic difference was confirmed by a perception experiment which showed that [a] and [ɛ] were more often confused with each other than [i] and [ɛ] – as well independent of nasalization (see chapter 2.3). Note that it was not checked whether these differences were statistically significant. Thus, [a] and [ɛ] as well as [ã] and [ɛ̃] are acoustically and perceptually more similar to each other than [ɛ] and [i] or [ɛ̃] and [ĩ] – in general and for German native speakers.

If this phonetic similarity is active during morphophonological learning, one will expect that learners, once acquired a nasalization pattern, will generalize it to the phonetically most similar pattern. I tested this by means of an artificial language learning experiment (see chapter 2.4) and I found a generalization advantage for the phonetically similar pattern. This means that there were more generalizations to [ɛ] than to [i] by [a]-learners, more generalizations to [a] than to [i] by [ɛ]-learners and more generalizations to [ɛ] than to [a] by [i]-learners. I further tested for a bias based on phonetic similarity by means of a productive test with French loan words (see chapter 3). In this experiment I found only a generalization advantage to the phonetically most similar pattern for [i]-learners, but not for [a]-learners or for [ɛ]-learners. However, the results were highly variable and none of the other factors – ease of perception or ease of articulation – predicted the generalization results correctly. Hence, phonetic similarity was the factor whose predictions were confirmed at least for one learning group.

Figure 31 illustrates how phonetic similarity affects generalizations of a learned vowel nasalization pattern. Note that for illustrative purposes it is assumed that phonetic similarity is the only factor affecting generalizations. Low [ã] and mid [ɛ̃] are phonetically more similar to each other than mid [ɛ̃] and high [ĩ] as shown by the different distances of the circles in the Figure. The more phonetically similar trained and untrained patterns are, the easier it is to generalize the trained pattern to the untrained pattern. This is illustrated with the thickness of the arrows: Thick arrows represent easy generalizations (from mid to low, and vice versa), thin arrows represent more difficult generalizations (from mid to high, and vice versa) and dotted arrows represent difficult generalizations (from high to low, and vice versa). This generalization disadvantage for high vowel nasalization is independent of the learning advantage for high vowel nasalization due to ease of perception, which is illustrated with a thick circle around high [ĩ] here. The reason for this independence is that no group failed to learn the trained pattern. All participants learned their trained pattern, however with different ease (see Figure 28).

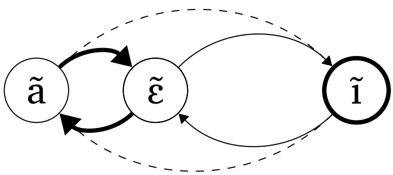

Figure 31: Generalizing vowel nasalization based on phonetic similarity.

To conclude, learners seem to be biased by phonetic similarity when generalizing a trained pattern to untrained items. This means that phonological learning is biased by phonetic similarity in such a way that it is easier to transmit a pattern to a phonetically similar pattern than to a phonetically more distant pattern. This in turn means, that phonetic details are stored in our mental lexicon which are actively used during the process of learning and during the process of generalizing phonological patterns.

My results are in line with Kapatsinski (2013a)'s interpretation of Wilson (2006)'s results which show that learners generalize to phonetically similar patterns. Learners were exposed to items in which [k] underwent velar palatalization either before [e] or before [i] and to items in which [k] did not change before [a]. At test, [e]-learners generalized velar palatalization to [i]-items, whereas [i]-learners did not generalize velar palatalization to [e]-items. As [e] is equally similar to [a] and to [i] it was not clear for [i]-learners based on phonetic similarity whether [e]-items have to underwent velar palatalization such as [i]-items or whether they do not have to change like [a]-items during exposure. Further support for a role of phonetic similarity comes from Cristià et al. (2013) who found that generalizations are made based on phonetic similarity between training items and test items.

One might assume that learners expand the phonetic category they were exposed to so that it covers phonetically similar patterns as well. This may be done to prevent possible misperceptions – especially because the pattern they are exposed to during training is a novel one they did not know (phonemically) from their native language. In this way the formed category of a certain nasal vowel may be greater than actually necessary in order to cover similar sounding patterns as well. Figure 32 illustrates such an expansion of the phonetic category, which would lead to overlapping categories for mid and low nasal vowels in this case.

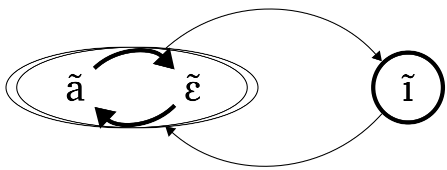

Figure 32: Expansion of the learned category due to phonetic similarity does not explain the generalization results.

However, the learners in my experiments did not make use of such an expansion of the learned category as this expansion would not explain why there are fewer generalizations from high to low than to mid. If participants had expanded their phonetic category to prevent misperceptions, I would have expected equal generalizations from high to mid and to low. This, however, was not found.

6.2.5 The role of perceptual similarity to L1-phonemes

I studied the role of perceptual similarity to native language phonemes in the experiments on retroflexion. In a perception experiment (see chapter 4.3) I found that the native contrast between alveolar and postalveolar fricatives as well as the contrast between alveolar and retroflex fricatives were perceived best, whereas the contrast between postalveolar and retroflex fricatives was perceived worst by German native speakers. These results suggest that German native speakers mapped unfamiliar, non-native retroflex fricatives onto familiar, native postalveolar fricatives, which means that they did not perceive the fricative as retroflex but as postalveolar. This is due to the perceptual similarity of the retroflex fricative to the native postalveolar fricative. German native speakers were also unsure about the status of the retroflex consonant as native phoneme – 33% of all participants identified it as native. The same is true for the postalveolar fricative which was identified by 30% of all participants as non-native. It is thus possible that both fricatives are perceived as one sound: as native postalveolar fricative.

If this perceptual similarity to native language phonemes is active during phonological learning, one will expect that it is easier to learn a pattern concerning sounds which can be mapped onto native sounds – in this case the fricative pattern – than a pattern which concerns non-native sounds which cannot be mapped onto native sounds – in this case the patterns with plosives and lateral approximants. Further, one will expect that learners, once acquired the fricative pattern, will not be able to generalize it to other consonant manners, as the retroflex lateral approximants and retroflex plosives do not have a perceptually similar postalveolar counterpart in German onto which the retroflex consonants can be mapped. Thus, learners exposed to fricatives will not be able to generalize the perceived alternation between alveolar and postalveolar fricatives on alveolar and retroflex plosives or lateral approximants. Consequently, learners, once acquired the retroflexion pattern with plosives or lateral approximants, will not be able to generalize it to fricatives, as they will perceive the retroflex fricative as postalveolar as well.

I tested this by means of artificial language learning experiments. In the artificial language learning experiment in which evidence for a morphophonological alternation was provided (see chapter 4.4) I found a learning advantage for fricatives which means a learning advantage for the pattern concerning sounds which can be mapped onto native sounds. In addition to that I found that learners exposed to fricatives did not generalize their learned pattern to other consonants and that learners exposed to either plosives or lateral approximants did not generalize their trained pattern to fricatives.

Figure 33 is an adaptation of Figure 29 and illustrates how the perceptual similarity to native language phonemes affects learning and generalizing retroflexion. Note that for illustrative purposes it is assumed that the perceptual similarity to native language phonemes is the only factor affecting learning and generalizing retroflexion. Alveolar and retroflex fricatives as well as alveolar and postalveolar fricatives are perceptually more distant from each other than alveolar and retroflex plosives or lateral approximants, which is illustrated with the different distances of the circles in the Figure. The thick circles represent the native sound categories – the alveolar and postalveolar consonants – and the solid and dotted circles represent the novel sound categories – the retroflex consonants. The non-native retroflex fricative and the native postalveolar fricative are perceptually very similar to each other (see the overlapping circles in the Figure) so that the retroflex fricative is perceived as native postalveolar (Flege, 1995; Lehiste, 1988). As alveolar and postalveolar fricatives contrast in German, this contrast is already known by the learners. This is why the alternation from alveolar fricatives to postalveolar fricatives is learned better than the alternation from alveolar lateral approximants or plosives to retroflex lateral approximants or plosives, which do not form a contrast in German as the retroflex consonants are not part of the German phoneme inventory (Wiese, 1996).

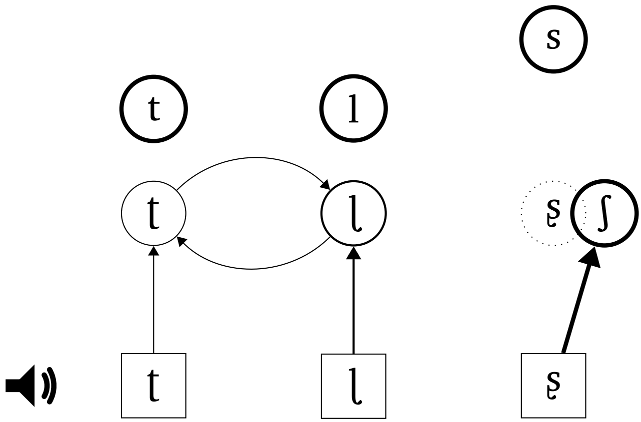

Figure 33: Learning and generalizing retroflexion based on perceptual similarity to native postalveolar fricative.

In this Figure arrows reflect the learning of the unfamiliar, non-native retroflex consonants (circles) and the mapping of the unfamiliar, non-native retroflex fricative onto the familiar, native postalveolar fricative by means of auditory exposure to retroflex consonants (rectangles). Thick arrows mean a better learning than thin arrows. Postalveolar fricatives are learned best, which is illustrated with the thick circle around them in comparison to the solid circles around retroflex lateral approximants and plosives. However, as exposure to retroflex fricatives does not end up in learning a retroflex fricative, but in learning a postalveolar fricative, there are no generalizations from postalveolar fricatives to retroflex consonants. Conversely, as learners of retroflex plosives and lateral approximants perceive the retroflex fricative as postalveolar as well, they cannot generalize their learned retroflexion pattern to fricatives. Nonetheless, learners of retroflex lateral approximants can generalize their learned retroflexion pattern to plosives, and vice versa, which is illustrated with the bended arrows connecting the learned categories in the circles.

However, Figure 33 does not fully explain the learning and generalization behavior of the participants. There is one more perceptual factor which affects learning and generalizations. Native Germans had difficulties identifying the retroflex [ʈ] because they confused it not only with its alveolar counterpart [t] (8%) but also with its voiced counterparts [d] (13%) and [ɖ] (52%). This is the reason why participants exposed to plosives performed worst, which means that they did not learn their retroflexion pattern and consequently did not generalize their pattern to other consonants. Figure 34 illustrated this.

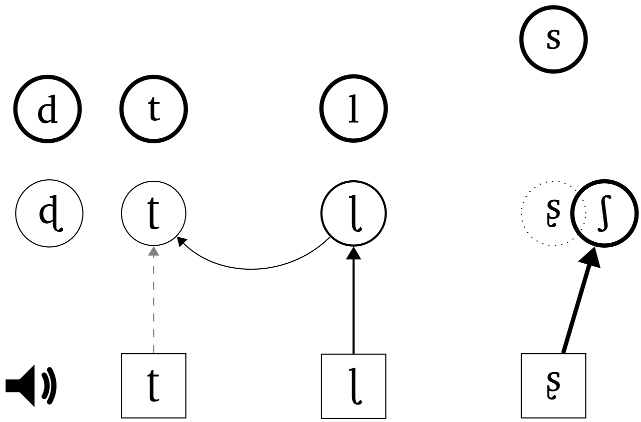

Figure 34: Corrected version of learning and generalizing retroflexion based on perceptual similarity to native postalveolar fricative.

For reasons of simplicity the distances between [ʈ], [t], [d] and [ɖ] are held constant, although [ɖ] is – based on the confusion data (see Table 15) – perceptually more similar to [ʈ] than the other plosives.

In the artificial language learning experiment in which no evidence for a morphophonological alternation was provided (see chapter 5) I found a learning advantage for fricatives as well. However, the predictions for the generalizations could only be confirmed for fricative learners and partly for plosive learners, which means that the perceptual similarity to native language phonemes biases learners more when evidence for an alternation is present and hence training is more reliable than when there is less evidence for an alternation and hence training is less reliable.

Note that ease of perception (see chapter 6.2.2) predicts the same outcome. Thus, it cannot be stated whether the learning advantage for retroflex fricatives is due to the perceptual similarity to native language phonemes, ease of perception or both of them.

The perceptual similarity to native language phonemes seems to bias phonological learning as well as generalizations to untrained patterns – especially in reliable training conditions. However, one has to keep in mind that a direct mapping from a non-native category onto a native one – caused by a very small perceptual distance between non-native and native category – may not only result in a learning advantage but also in mislearning: Generalizing a learned pattern, which was mapped onto a native one, to other consonants is hampered by a direct mapping onto a native category because the learners did not perceive and learn what they were supposed to perceive and learn. During learning it does not matter whether one or both of the involved sounds of a contrast are incorrectly mapped onto a native category. However, when generalizing the learned contrast to other sounds, it is important how the involved sounds are mapped. An incorrect mapping onto a native category will thus result in an incorrect learned rule which cannot be generalized correctly to other sounds.

It is further important to stress that the influence of perceptual similarity to native language phonemes on learning vowel nasalization could not be tested, as the most perceptually similar native phonemes to the nasal vowels are their oral counterparts. A mapping of nasal vowels onto their oral counterparts, however, would not results in any alternation so that is would be impossible to learn a phonological rule or pattern.

My findings are in line with the assumptions made by Best (1995), Flege (1995) and Kuhl & Iverson (1995) insofar as non-native sounds are more likely to be perceived as native, the less perceptually distant the non-native sounds are to a native category. Concerning learning non-native sounds, Flege (1995) and Kuhl & Iverson (1995) claimed that it is easier to learn a non-native sound which is perceptually distant from a native category, whereas Aoyama et al. (2004) claimed that it is easier to learn a non-native sound which is perceptually similar to a native category. I agree with both of them. Aoyama et al. (2004)'s claim reflects the pure learning results of my participants: Retroflex fricatives were learned best because they are perceptually similar to native postalveolar fricatives. Flege (1995)'s and Kuhl & Iverson (1995)'s claim reflects the generalization results of my participants – learning in terms of applying a learned pattern to novel patterns: Fricative learners did not generalize, whereas lateral approximant learners generalized their learned retroflexion pattern to plosives.

6.2.6 The role of structural simplicity

Although I tried not to confound phonetic naturalness with structural simplicity, I may have found an effect of structural simplicity in phonological learning in the experiments on retroflexion. In the perception experiment on retroflexion (see chapter 4.2) I found that in plosives not only retroflexion was often misperceived but also voicing. Hence, in the artificial language learning experiments on retroflexion (see chapters 4.4 and 5) the poor results of the [t]-learners may not only be due to their perceptual difficulty to identify the retroflex voiceless plosive correctly as elaborated above and illustrated in Figure 34 but also due to their perceptual difficulty to perceive voicing correctly. This misperception of voicing might have lead to a perceptually more complex alternation for [ʈ]-learners compared to those perceived by [l]-learners and [s]-learners in training. Thus, based on the perceptual confusions I may give evidence for a learning advantage for structurally simple patterns (one-feature change) over structurally complex patterns (two-feature change). However, as the presented items were recorded as containing [t] and [ʈ] and not as containing [t] and [ɖ], they objectively do not contain a two-feature change. The differences in complexity hence likely evolve by the misperception of [ʈ] as [ɖ] by the participants. If this is really the case in the experiments, will nevertheless remain speculative.

Nonetheless, there are numerous studies giving evidence for a simplicity bias in phonological learning, e.g. Glewwe (2019) and Moreton (2008, 2012). In addition to that there are further studies having confounded phonetic naturalness with simplicity, thus giving evidence for a simplicity bias as well (Peperkamp et al., 2006; Pycha et al., 2003; Skoruppa & Peperkamp, 2011).

6.3 Factors affecting the emergence of a phonetic bias

I will now describe which factors influence the emergence of a phonetic bias in artificial language learning experiments based on my findings. These factors are on the one hand the kind of test – productive test or perceptual test – and on the other hand the kind of training – providing evidence for a morphophonological alternation or not. Doing so, I will evaluate if and how these factors influence the outcome of a morphophonological learning experiment and thus, why a certain phonetic bias was found.

6.3.1 Perceptual test vs. productive test

In the experiments on vowel nasalization I studied the effect of different kinds of tests on the emergence of a phonetic bias. As ease of perception and ease of articulation make different predictions regarding the learning and generalization advantage of nasalization of different vowel heights, it may be the case that both predictions can be confirmed depending on whether participants are tested on their articulation or on their perception of vowel nasalization. To briefly recapitulate: The articulation of low vowel nasalization is easiest and the perception of high vowel nasalization is easiest. This is true for German native speakers as well, as a production experiment (see chapter 2.2) and a perception experiment (see chapter 2.3) showed.

In the artificial language learning experiment with a perceptual test (see chapter 2.4) I found a learning advantage for the perceptually easy pattern, but not for the articulatory easy pattern. This means that participants who had been exposed to high vowel nasalization performed better on the trained pattern at test than participants who had been exposed to mid or low vowel nasalization. A similar result was observed in the recordings of the French loan words before and after exposure to vowel nasalization (see chapter 3). Participants who had been exposed to high vowel nasalization did not modify the amount of nasality after exposure compared to the amount of nasality before exposure, whereas participants exposed to mid and low vowel nasalization decreased their amount of nasality after exposure, which indirectly increased the amount of nasality after exposure to high vowel nasalization. This means that participants performed better with the perceptually easy pattern in a productive test as well as in a perceptual test.

Whether the kind of task influences the generalization behavior instead, cannot be decided based on these experiments. In both experiments phonetic similarity explained the generalization behavior of the participants better than ease of perception or ease of articulation did. However, a possible effect on the generalization behavior might be figured out with an experiment in which no other factors, e.g. phonetic similarity, bias learning and generalizations. It might then be possible that the generalization behavior changes depending on the kind of task.

To conclude, the kind of task – productive test or perceptual test – does not seem to have an effect on whether a bias based on perception or a bias based on articulation affects phonological learning. Independent of the task a perceptual bias affects learning, whereas an articulatory bias does not affect learning. This is in line with Glewwe (2019) who claimed that a bias based on perception shows up more often than a bias based on articulation, as well as with Flege (1991)'s perception before production hypothesis. According to this hypothesis "perception leads production" (Flege (1991), page 264), which means that learners need to accurately perceive a non-native sound before they are able to accurately produce it. Conversely, it rarely, if ever, appears that learners accurately produce a non-native sound without being able to perceive this sound. Moreover, due to the variable and weak results in the experiment with the productive test (see chapter 3) it seems as if generalizing a perceived pattern to production is difficult and that even a small modification of already known pronunciations depends on many factors, one of which is phonetics.

6.3.2 Evidence for a morphophonological alternation

In the experiments on retroflexion I studied the effect of different kinds of trainings, which either contained evidence for a morphophonological alternation or not, on the ability to generalize in general and on the emergence of a phonetic bias. As previous studies on the one hand have shown a transfer of a phonological feature from one form to novel forms without evidence for an alternation (Berent et al., 2001, 2002, 2007) and on the other hand a greater influence of phonetics than of L1 in unreliable trainings, e.g. short training phases or training sets with exceptions, and a greater influence of L1 than of phonetics in reliable trainings, e.g. long training phases or training sets without exceptions (Baer-Henney et al., 2015a; Greenwood, 2016), I conducted two artificial language learning experiments on retroflexion which differed only in whether they provide evidence for a morphophonological alternation or not – hence in how reliable their training is. The first experiment (see chapter 4.4) included singular, plural and diminutive forms. This was done to ensure that participants learn that the consonant is only retroflexed in plurals and in the context of [ɾ] – thus providing evidence for an alternation. The second experiment (see chapter 5), however, included plural and diminutive forms only. This reduced the evidence for a morphophonological alternation and made the training in general shorter than in the first experiment.

In the artificial language learning experiment with evidence for an alternation during training at least one group of learners generalized the trained pattern to novel items, whereas in the artificial language learning experiment without evidence for an alternation during training no group generalized the trained pattern to novel items. This speaks against the experimental results which were explained by pure transfer mentioned above. Nevertheless, I found more evidence for a bias based on the perceptual similarity to native language phonemes, thus for an influence of L1, in the artificial language learning experiment with evidence for an alternation during training than in the artificial language learning experiment without evidence for an alternation during training. The performance of the participants to untrained items in the latter, however, seems to be affected by ease of perception, thus by phonetics. This is in line with the previous studies mentioned above.

It thus seems as if learners need knowledge about the relation of forms to accurately predict the phonological form of novel items based on familiar items. This underlines the importance of the relation between forms in a paradigm to accurately acquire a phonological pattern – including the ability to generalize this phonological pattern to novel items. Without a relation between forms generalizations are not possible. Although this is not in line with Berent (2013) and Berent et al. (2012), one has to keep in mind that they tested generalizations of identity restrictions, e.g. AAB vs. ABB. The experiments on retroflexion, however, do not deal with patterns showing identity restrictions but with patterns showing alternations. Hence, Berent (2013)'s and Berent et al. (2012)'s assumptions may hold for generalizations of identity restrictions but – as the present experiment on retroflexion showed – not for generalizations of alternating patterns.

Additionally, the need for the evidence of a morphophonological alternation gives hints at the importance of allophones during speech perception. Whereas a generalization in case of no evidence for an alternation would have indicated that the learners interpreted the alveolar and the retroflex consonants under investigation as two different phonemes, a generalization only in case of evidence for an alternation would have indicated that the learners interpreted the alveolar and the retroflex consonants under investigation as allophones of one phoneme. The failure of the learners to generalize the retroflexion pattern without having evidence for a phonological alternation can thus be taken as hint that allophones are used in the process of speech perception – and learning – as well. This is in line with the research on perceptual learning by Mitterer & Reinisch (2017), Mitterer et al. (2013, 2018) and Reinisch et al. (2014) – although in the present experiments contrary to those investigating perceptual learning no lexical access was required.

It further seems as if learners rely on phonetics – especially on perceptual ease – first or in situations in which it not clear if and how the morphophonological forms are related to each other when learning phonology. This can be transmitted to young children who are also biased more by phonetics than by their L1 as they do not know every L1-characteristic properly yet (van de Vijver & Baer-Henney, 2014). However, as soon as another (more) reliable source is available to guide learning, learners downgrade the reliance of perceptual ease and more strongly rely on phonetic, phonological or lexical characteristics of their native language. However, if there is no support from L1, even adults are biased by phonetics (Baer-Henney &

van de Vijver, 2012; Finley, 2012; Wilson, 2006).

These effects of phonetics and the native language dependent on whether evidence for a morphophonological alternation is provided during training or not is illustrated in Figure 35.

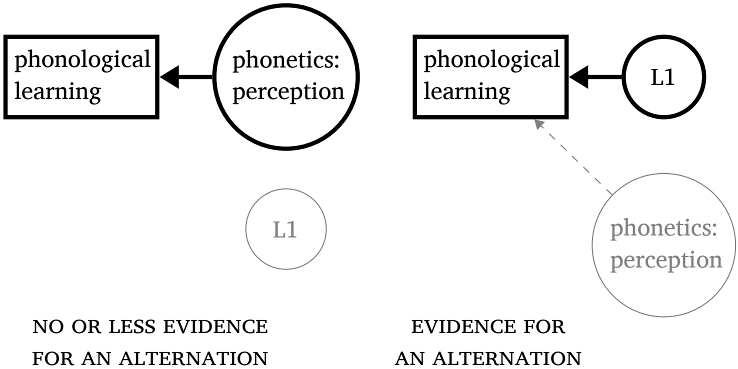

Figure 35: Effects of phonetics and L1 on phonological learning dependent on evidence for a morphophonological alternation.

In the left part you can see that with less or no evidence for an alternation phonological learning is mainly affected by phonetics, e.g. perceptual ease, which is illustrated with the black circle and the thick arrow showing the influence of phonetics on phonological learning. The L1, e.g. phonotactics, does not affect phonological learning as illustrated with the gray circle and the missing connection to phonological learning. In the right part of the Figure you can see that the more evidence for an alternation is provided during training or the more experience learners gain, the weaker the influence of phonetics becomes and the stronger the influence of the L1 which downgrades phonetics becomes. This is illustrated with the upwards movement of L1 and the downwards movement of phonetics compared to the illustration in the left part of the Figure. Besides, L1 is now connected by a black and thick arrow to phonological learning, whereas, the influence of phonetics on phonological learning is weaker which is illustrated with the dashed arrow in gray.

In the case of retroflexion, the different effects of phonetics and L1 dependent on the evidence for a morphophonological alternation during training specifically mean that the mapping from the retroflex fricative changes (see Figure 36). In the training condition without evidence for a morphophonological alternation learners rely on ease of perception. What counts is the difference between the alternating sounds – alveolar and retroflex fricative – illustrated with the distance between the circles in the Figure. In this case the retroflex fricative would be correctly mapped on a retroflex fricative – illustrated with the arrow connecting the rectangle (exposure sound) and the thick circle (learned category) with the retroflex fricative in the left part of the Figure. However, in the training condition with evidence for a morphophonological alternation learners are not biased by ease of perception that heavily anymore. L1 is used as a source for reliable information as well, and as this is available the influence of ease of perception becomes weaker – illustrated with the dotted circle around the retroflex fricative in the right part of the Figure –, whereas the influence of L1 becomes stronger. Consequently, the retroflex fricative is incorrectly mapped on the perceptually similar postalveolar fricative – illustrated with the arrow connecting the rectangle (exposure sound) and the thick circle (learned category) with the postalveolar fricative in the right part of the Figure. In this case, it does not matter that retroflex and alveolar fricatives are perceptually distant from each other or not. What counts is that postalveolar (and alveolar) fricatives are native consonants. Dependent on whether retroflex fricatives are perceived as retroflex or as postalveolar fricatives, there are either generalizations from fricative learners to the other consonants and generalizations from the other learners to the retroflex fricatives or there are no generalizations from fricative learners to the other consonants and no generalizations from the other learners to the postalveolar fricative. For reasons of simplicity the different generalizations are not shown in Figure 36.

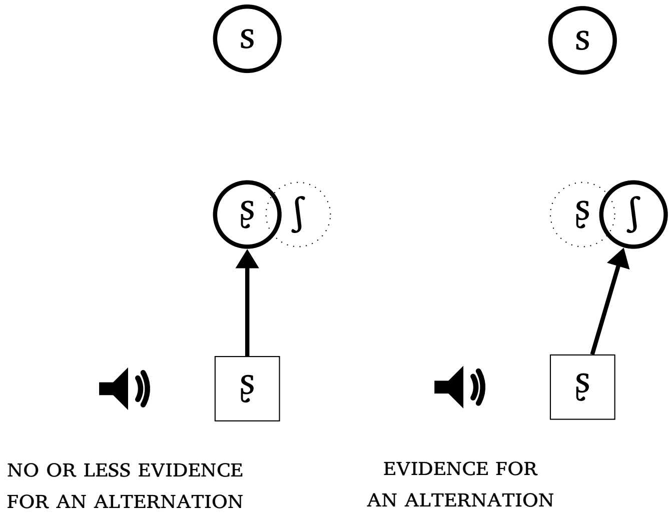

Figure 36: Effects of ease of perception and perceptual similarity to L1 on learning retroflexion dependent on evidence for a morphophonological alternation.

6.4 Proposed effects of a phonetic bias

Based on the experimental results and the evaluation of the phonetic factors as well as of factors influencing the emergence of a phonetic bias I will now propose which effects a phonetic bias might have. Doing so, I will argue that a phonetic bias does not seem to affect the cross-linguistic distribution of vowel nasalization as well as of retroflexion. However, synchronic phonology is biased by phonetic factors, most of which are perceptually motivated. This phonetic bias affects phonological learning and guides learners during the acquisition of an unfamiliar pattern as well as during the generalization of this pattern to other patterns. Learning and generalizations are in this case independent from each other and can be affected by different phonetic factors.

6.4.1 Effects of a phonetic bias on markedness

According to the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006) the predispositions based on ease of articulation and ease of perception are encoded as markedness in our grammars and thus are the reason for the typological distributions of certain sound pattern in the languages of the world: Typologically common patterns are those that are preferred by the learner over typologically rare patterns. I compared the typological distribution of each pattern with its ease of articulation and its ease of perception as well as with the learning results for each pattern not only for vowel nasalization but also for retroflexion. If phonetics affects typology in the proposed way, one will assume that the effects of ease of articulation and ease of perception lead to asymmetric distributions of sound patterns. For vowel nasalization this means more low and high nasal vowels compared to mid nasal vowels in the languages of the world. As low nasal vowels are preferred articulatory (Bell-Berti, 1993; Ohala, 1975) and as high nasal vowels are preferred perceptually (Schwartz, 1968), high and low nasal vowels should be used more often cross-linguistically than mid nasal vowels, which are neither preferred articulatory nor perceptually. However, nasal vowels of all three heights are equally distributed across the languages of the world (see chapter 2.1). Thus, both factors do not seem to be (the only) factors responsible for this equal distribution of nasal vowel height. Note that I only considered phonemic nasal vowels in the typological survey. It may be the case that allophonically nasalized vowels will lead to a different typological distribution, which might be in line with the predictions based on ease of perception and ease of articulation. A similar picture emerges in the case of retroflexion when comparing the typological distribution with its ease of perception and its ease of articulation as well as with the learning results. Retroflex fricatives are preferred perceptually (Kohler, 1990), whereas for ease of articulation the predictions differ: Either no difference between consonant manners is assumed (Stausland Johnsen, 2012) or retroflex fricatives are dispreferred articulatory (Hamann, 2003b). However, plosives are the most common retroflex consonants cross-linguistically, followed by fricatives and lateral approximants. Nasals are the least common retroflex consonants (of all consonants under investigation in the experiments described here) (see chapter 4.1). Thus, neither ease of articulation nor ease of perception can explain the asymmetric distribution of retroflex consonants across languages in the world.

Proponents of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006) further claimed that learning and generalizations are affected by ease of perception and by ease of articulation. I can confirm this for learning but not for generalizations and I can confirm this only for ease of perception and not for ease of articulation based on the vowel nasalization results (see Table 6). Concerning retroflexion, I can say that learning may be biased by ease of perception as well. However, as L1 predicts the same outcome, it cannot be said for sure that the learning results are (solely) due to ease of perception. Learners' generalizations of retroflexion are biased by ease of perception but only when training provides no evidence for an alternation and is thus unreliable (see Table 25). With evidence for an alternation and thus with a more reliable training, ease of perception does not affect generalizations (see Table 20). Thus, my findings provide only limited evidence for a learning bias as a result of phonetic markedness.

In addition to that the German participants benefit from ease of perception during the identification of nasal vowels (see chapter 2.3) and retroflex consonants (see chapter 4.2) and they benefit from ease of articulation during the production of nasal vowels (see chapter 2.2). German native speakers are thus guided by phonetics when producing or perceiving unfamiliar sound patterns as for example nasal vowels of different height or retroflex consonants. This is in line with the framework of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006). Note that the learning advantage for retroflex fricatives may also be due to the perceptual similarity to the German postalveolar fricative.

In sum, the results of the artificial language learning experiments show that learning and especially generalizing morphophonological patterns are not necessarily guided by markedness. However, as sound patterns in languages are the result of change in hundreds of years, this change is affected by many factors only one of which is the phonetic grounding of a sound pattern. Hence, not all learning biases directly affect typological patterns. Further, it is important to stress, that there are phonetic factors that influence synchronic morphophonological generalizations (see chapter 6.4.2 for details). The vowel nasalization learners generalized from trained patterns to untrained patterns in a way that reflects the phonetic similarity between the trained patterns and the untrained patterns. This is important because it gives evidence for a role of phonetics in synchronic phonology, which is against the assumptions of Blevins (2004), Hale & Reiss (2000) and Ohala (1986). Blevins (2004) argued against a learning bias and assumed that phonetics shapes phonology only diachronically. The contrastive use of acoustically and perceptually similar sounds hampers effective communication. To overcome this difficulty the contrast between these sounds disappears gradually over time. Experimental results, however, showed that learners do use phonetics to make phonological synchronic generalizations (White, 2017; Wilson, 2006; Zuraw, 2007). The proposal that all phonetic explanation in phonology is diachronic is therefore too strong, and learners clearly make use of phonetic similarity to make phonological generalizations as my results on vowel nasalization show. Independent of this, ease of perception and ease of articulation may affect the phonology of languages diachronically. As sound patterns in languages were – and still are – affected by many factors not all phonetic learning biases directly affect typological distributions.

6.4.2 Effects of a phonetic bias on synchronic phonological learning and generalizations

In the previous chapter it was already mentioned that phonetics affects how morphophonological patterns are learned and generalized. I investigated the role of phonetics in learning and generalizations not only in the artificial language learning experiments on vowel nasalization (see chapters 2.4 and 3) but also in the artificial language learning experiments on retroflexion (see chapters 4.4 and 5). Learning refers to how well participants performed in the trained patterns, whereas generalizations refer to the learners' performance in the untrained patterns.

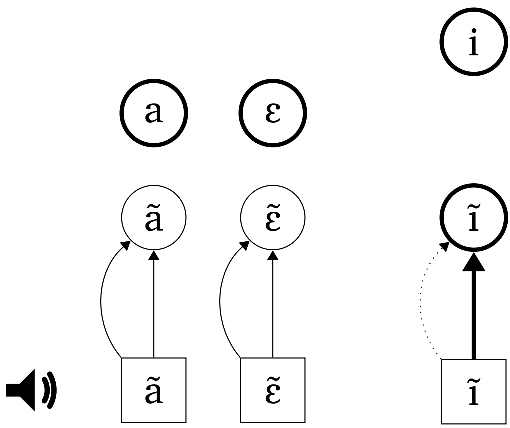

Figure 37: Learning vowel nasalization.

In the learning experiments on vowel nasalization it was tested whether the learning behavior is biased by ease of articulation, ease of perception or perceptual frequency (the latter only in the experiment in chapter 3). It turned out that the learning behavior could be explained best with a learning bias based on ease of perception. A bias based on perceptual frequency can explain the results in the experiment described in chapter 3 as well, as the predictions based on ease of perception and perceptual frequency do not differ in this case. Figure 37 is a combination of Figure 28 and Figure 30. It illustrates how learning is biased by ease of perception and perceptual frequency. The more distant oral and nasal vowels are and the weaker the memory traces (bended arrows) of the nasal vowels are, the better the nasal vowels are learned – illustrated with the thick arrow and the thick circle around the high nasal vowel.

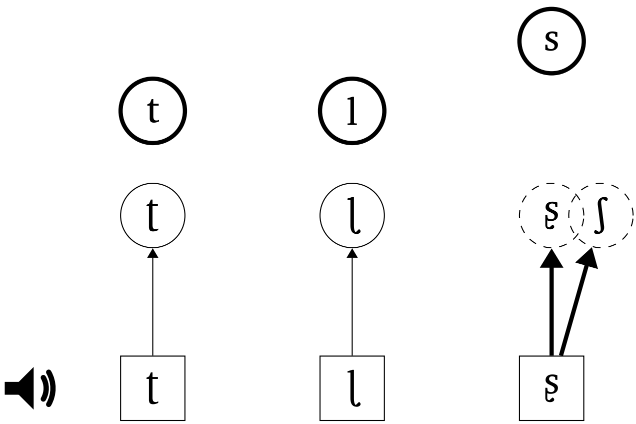

Figure 38: Learning retroflexion.

In the artificial language learning experiments on retroflexion it was tested whether the learning behavior is biased by ease of perception or by the L1 caused by the perceptual similarity to native German phonemes. It turned out that the learning behavior could be explained not only with a learning bias based on ease of perception but also with a learning bias based on L1, as both make the same predictions. Figure 38 is a combination of Figure 29 and Figure 33. It illustrates how learning is biased by ease of perception and the perceptual similarity to the German postalveolar retroflex. The more distant alveolar and retroflex consonants are and the closer a perceptually similar native sound is, the better the retroflex consonant is learned.

Unfortunately, ease of perception and perceptual frequency as well as ease of perception and the perceptual similarity to L1-sounds make the same predictions for learning. It is thus impossible to tease apart these factors with the present data to find out which of them is the one that affects learning or which of them affects learning more. To investigate which of the factors is given more weight, it would be necessary to test the learnability of a pattern for which the predictions differ depending on the three factors.

Concerning the generalization behavior, the learning experiments on vowel nasalization investigated a bias based on ease of articulation, ease of perception and phonetic similarity. The results show that a bias based on phonetic similarity explains the generalization results best. For an illustration see Figure 31. The learning experiments on retroflexion investigated the influence of ease of perception and of the L1 caused by the perceptual similarity to native language phonemes. It turned out that the emergence of a bias depends on whether evidence for a morphophonological alternation is provided during training or not: In a training condition in which evidence for an alternation is provided the perceptual similarity to native language phonemes is given more weight, whereas in a training condition in which no evidence for an alternation is provided ease of perception seems to affect the performance in the untrained conditions more (see Figure 35). Note that the lack of evidence for a morphophonological alternation actually prevents generalizations. However, whether the lack of evidence for an alternation really prevents generalizations or whether it just hampers them cannot be stated with the present experiments as the interpretation mainly relies on the data of one learning group ([l]-learners) which only made up about 20-25 participants in each experiment. Further research on the importance of evidence for a morphophonological alternation during phonological learning is thus necessary.

Putting the results of the artificial language learning experiments on vowel nasalization and retroflexion together, it seems as if the learning behavior is mainly biased by ease of perception (and/or by perceptual frequency). In the case of retroflexion it may also be biased by the perceptual similarity to native language phonemes, thus by L1. However, as the predictions based on perceptual similarity to native language phonemes are the same as those based on ease of perception, this is not entirely clear. It further seems as if learners form generalizations independent of what is easy to perceive or articulate, but based on the perceptual or phonetic similarity to the trained pattern or to native language phonemes, thus, to familiar sounds. The only exception to this is found when no evidence for an alternation is provided during training. In this case ease of perception seems to bias the performance of the learners. Whether the performance in the untrained conditions would be biased by ease of perception in a vowel nasalization experiment without evidence for an alternation needs to be tested to make a clear statement about this. To test this, the artificial language learning experiment on vowel nasalization (see chapter 2.4) needs to be run without singular forms in the training and at test. If participants would then make generalizations based on ease of perception, the lack of evidence for a morphophonological alternation may be the reason why a bias based on ease of perception emerges. Note, however, that the lack of evidence for an alternation might also cause a lack of generalizations as the experiment on retroflexion has shown (see chapter 5).

The following Figure 39 illustrates how each of the phonetic biases affecting learning and generalizations works. Perceptual frequency is left out here for reasons of simplicity and because perceptual frequency does not make any predictions concerning generalizations. In this Figure it is shown how different phonetic (or perceptual) distances between sounds affect the learning and the generalization behavior of learners. I will focus on the perceptual distances between alternating sounds (X and X'), which is the basis for ease of perception, on the perceptual distances between a non-native sound and the perceptually most similar native sound (X and X"), which is the basis for the perceptual similarity to native language phonemes, and on the phonetic distance between trained patterns (A, B and C) which is the basis for phonetic similarity. Note that even phonetic similarity includes perceptual similarity, as phonetic similarity is mainly determined by means of perceptual, acoustic and featural similarities.

In Figure 39 the effects of phonetic biases are shown separately. Each bias is illustrated with three different patterns (A, B and C) which either do not differ depending on phonetics (top) or which differ in their phonetic similarity (second from top), in their perceptual similarity to native language phonemes (third from top) or in their ease of perception (bottom). X refers to the unfamiliar, non-native sound, X' refers to the familiar, native sound and X" refers to the familiar, native sound which is perceptually most similar to X.

In the illustration on top of Figure 39 none of the described distances differs from each other. There is the same distance between X and X' and between X and X" for the sound patterns A, B and C. Further, A, B and C are equally distant from each other. In this case one expects neither a learning nor a generalization advantage for one of the patterns.

The illustration below (second from above) shows a phonetic bias based on phonetic similarity. Here only the distances between A, B and C differ as well as those between A', B' and C' and A", B" and C", respectively – as illustrated with the gray circles. B is phonetically more similar to C than to A. A and C are perceptually least similar to each other. Note that there are no differences in the distances between X and X' which mean no differences in learning. The more phonetically similar X and Y are, the easier it is to generalize the pattern to each other.

The illustration below (third from above) shows a phonetic bias based on the perceptual similarity to native language phonemes. Here only the distances between X and X" differ – as illustrated with the gray circles. A and A" are perceptually most similar to each other – they even overlap –, followed by B and B". C and C" are perceptually most distant from each other. The more perceptually similar X and X" are, the easier it is to map X onto native X" and the easier it is to learn the pattern between X' and X (or X")[12]. However, one has to keep in mind that the mapping onto X" means mislearning: The more perceptually similar X and X" are, the more difficult it is to generalize to this pattern from other patterns and the more difficult it is to generalize to other patterns from this pattern. It is important to note that although the distance between X and X" can be gradient, the mapping from X onto X" is categorical: Either X and X" are perceptually similar enough to map X onto X" (e.g. A and A") or X and X" are perceptually too distant from each other to map X onto X" (e.g. B and B" or C and C").

The illustration at the bottom of Figure 39 shows a phonetic bias based on ease of perception. Here only the distances between X and X' differ – as illustrated with the gray circles. A and A' are perceptually most distant from each other, followed by B and B'. C and C' are perceptually most similar to each other. The more distant X and X' are, the easier it is to learn the pattern between X and X' and the easier it is to generalize to this pattern from other patterns. It is important to note that the learning advantage and the generalization advantage for perceptually easy patterns are independent from each other. It may be the case that learning is biased by ease of perception but that generalizations are not biased by ease of perception (see experiments on vowel nasalization). It is further important to note that ease of perception seems to affect generalizations only in unreliable training conditions.

[12]If there is more than one candidate for X", the candidate that is most similar to X is chosen.

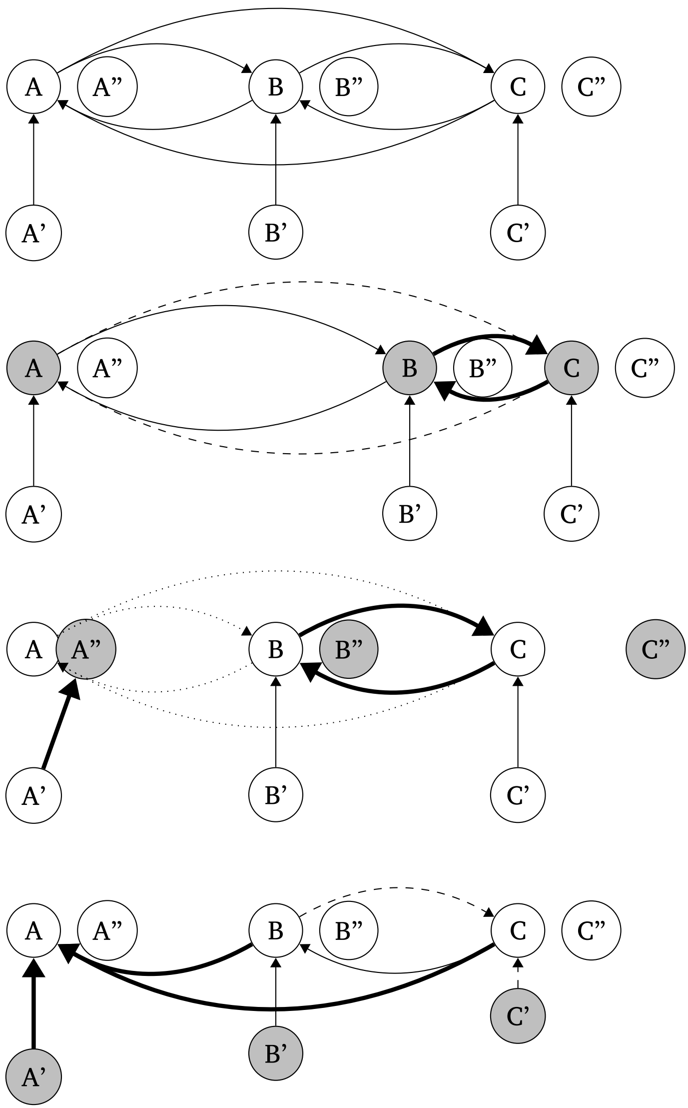

Figure 39: Phonetic biases affecting learning and generalizations.

To sum up, phonetic biases affect the learning behavior and the generalization behavior of learners exposed to a morphophonological pattern. The learning behavior is mainly biased by ease of perception. The perceptual similarity to native language phonemes – if available – guides learners in learning unfamiliar sound patterns as well. I did not find any evidence for a bias based on ease of articulation – not even in a productive task –, which means that a bias based on ease of perception is stronger than a bias based on ease of articulation, as has been shown by the experimental results on vowel nasalization. The generalization behavior is mainly biased by phonetic similarity between sound patterns or by the perceptual similarity to native language phonemes – if available. Ease of perception biases generalizations only when evidence for an alternation is missing during training. A bias based on phonetic similarity – if available – is stronger than a bias based on ease of perception, as has been shown by the experimental results on vowel nasalization. Further, a bias based on the perceptual similarity to L1-phonemes – if available – is stronger than a bias based on ease of perception as the experimental results on retroflexion have shown. The same might be true in the learning behavior as well.

Both mechanisms – learning and generalizations – are independent of each other. If learners are biased by ease of perception, they do not necessarily have to be biased by ease of perception when generalizing, and vice versa (see experiments on vowel nasalization). This suggests that several phonetic factors are pre-activated during phonological learning but are actively used at different stages of learning. It might be possible that ease of perception is more strongly activated at early stages of phonological learning but that its activation level falls as soon as other phonetic factors increase their activation level. At the beginning of the learning process – during training – learners are confronted with one sound pattern only. This pattern is either easy to perceive or difficult to perceive and it either resembles a native sound or not. Thus, only ease of perception and the perceptual similarity to native language phonemes can guide learners at this stage. However, when generalizing the newly learned sound pattern – at test –, other sound patterns appear. With these novel patterns new dimensions are present: Not only in the learned pattern can the distance between alternating sounds and the distance to native phonemes vary but also in the other patterns to which the learned pattern should be generalized. Further, the unfamiliar patterns can vary in their phonetic distance to the learned pattern. At this stage, the pre-activated factors phonetic similarity and perceptual similarity to L1-sounds downgrade ease of perception. Only if one or both of the similarity factors are not available because there are no similar sounds in the L1 or because the novel patterns are not of unequal phonetic similarity, thus if the similarity factors have no ability to guide learners during generalizations, then ease of perception will still be the most activated factor which then biases the learners' generalizations.

I suggest that the phonetic effects influencing phonological learning are best captured by means of contrasts or distances. In learning, those contrasts are learned best that contain the greatest perceptual distance. In this case it does not matter onto which kind of category (native or non-native) both sounds are mapped – what counts is that both sounds are perceptually far away from each other and thus mapped onto two different categories. Those different categories can either be both native, both non-native or one native and one non-native. In generalizations, the learned contrast is best generalized to another contrast that has the smallest distance to the learned one or to one of the contrasts that are learned best or better than the learned one.

Similar effects have been found in natural language acquisition studies. Stamer & Vitevitch (2012) found that advanced native English learners of Spanish were better at learning advanced Spanish words that sounded similar to known Spanish words (dense neighborhood) than words which sounded less similar to known Spanish words (sparse neighborhood). The learning performance was measured by means of a picture-naming task, a referent identification task and a perceptual identification task, and the similarity was measured by means of the neighborhood density, a measurement that takes into account how many phonologically similar words in the lexicon exist (Luce & Pisoni, 1998). This study differs in several aspects from the artificial language learning experiments on vowel nasalization: First, in the artificial language learning experiments participants were not asked to learn the lexical meaning of the words but were only tested on how well they learned and generalized the morphophonological alternation. Second, the neighborhood density to existing German words was not considered in the artificial language learning experiments on vowel nasalization so that any frequency counts are lacking. Third, the artificial language learning experiments did not make use of existing words but tested the learnability of words in a miniature language.

However, although the aims and the design of Stamer & Vitevitch (2012)'s study and of my experiments differ, both measurements – phonetic similarity and neighborhood density – are based on how similar different items sound. Hence, one can say that not only learning morphophonological alternations but also learning the meaning of words seems to be affected by the phonetic and phonological similarity between words: How well we learn novel words in a language depends – at least partly – on how similar they sound to already known words in this language – may it be an existing language or an artificial language. Stamer & Vitevitch (2012) refer to Storkel, Armbrüster & Hogan (2006) who suggest that representations established for a novel word are strengthened by phonological similarities to existing lexical representations (Jusczyk, Luce & Charles-Luce, 1994). In the case of vowel nasalization, however, the lexical representation is not accessed. Thus, phonological similarity seems to have an effect on the phonological representations as well, without any connection to lexical representations.

Similar observations have been made by Tang & Baer-Henney (under review) in an artificial language learning experiment with native speakers of German and Mandarin. They found that the neighborhood density of an item in an artificial language affects the participants' performance: The more similar an item was to another item in the artificial language – not in the native language – (dense neighborhood), the more often this item was termed as belonging to the artificial language the participants had been exposed to. This similarity effect is similar to the one found in the generalization behavior of my learners. It is true that I did not consider how many similar items there are and in addition to that I only considered the similarity between nasal vowels and not between the whole items, but independent of these differences an effect of phonological and phonetic similarity is observed which positively influenced the generalization behavior of the participants in both studies.

In chapter 1.1.2 I referred to Steriade (2001b, 2009) P-map theory to give evidence for a role of phonetic similarity in learning and generalizations. Based on this theory learners prefer perceptually minimal changes and generalize from dissimilar patterns to similar patterns more than vice versa. However, my results are not in line with the propositions based on the P-map. According to the P-map those alternations are learned best and most generalized to which constitute the smallest perceptual changes. Hence, in the case of vowel nasalization there should be a learning advantage and a generalization advantage for the smallest perceptual change between oral and nasal low vowels. I, however, found that those alternations were learned best which constitute the greatest perceptual change. Concerning the generalizations, I found an effect of phonetic similarity. However, this effect is different from the one proposed by the P-map. I did not observe more generalizations to the alternation between the most perceptually similar sounds but more generalizations to the alternation that was most phonetically similar to the learned alternation. This phonetic similarity depends also on perceptual factors that are due to acoustic similarity. To conclude, I cannot confirm the propositions based on the P-map theory. One reason may be that Steriade (2001b, 2009) based her assumptions on typological evidence – similar to the theory of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006). As the learning results on vowel nasalization do not conform to the typological distribution of nasal vowel height they may not confirm the propositions of the P-map theory either. Again, this might be due to the differentiation of contrastive vowel nasalization and allophonic vowel nasalization as pointed out in chapter 1.1.2.

7 Conclusion

The present dissertation aimed at investigating the role of phonetics in phonological learning by exploring the learning and generalization behavior of German native speakers when being exposed to either a vowel nasalization pattern or a retroflexion pattern. The main aim of this dissertation was to study the interaction of several learning biases all of which are grounded in phonetics. While previous studies have found evidence for the existence of such phonetic biases, their interaction and their contribution to the learning behavior on the one hand and their generalization behavior on the other hand, as well as their task dependency are yet unclear (see chapter 1.1). Exploring the interaction of these phonetic biases and the factors on which their emergence depends is, however, very important and necessary to accurately understand the role of phonetics in phonological learning.

As laid out in chapter 1.1, the experimental studies on phonological learning vary in their interpretations of their results concerning the role of phonetics. While proponents of phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006) argue that learners are predisposed towards phonetically motivated patterns and that this predisposition is encoded as markedness and reflected typologically, others, e.g. Kapatsinski (2013a) and Moreton & Pater (2012a,b), argue that these experimental results lack evidence for the existence of a phonetic bias by either explaining them with plain phonetic similarity or by pointing at confounding variables, e.g. L1-specific factors or structural simplicity, leading to inconclusive results. Some experimental results revealed effects of ease of perception, others of ease of articulation, phonetic similarity or the perceptual similarity to native sound categories. However, the experimental studies differ in whether they tested the production or the perception and whether they tested the learning or the generalization behavior. These differences in the experimental designs may have influenced which kind of phonetic bias, if at all, showed up and may thus have led to different interpretations and varying support for the existence of a phonetic bias. Nevertheless, these studies showed that phonological learning is not completely independent of phonetics.

To further investigate possible effects of phonetics on phonological learning and their interaction as well as their contribution dependent on type of test, evidence for a morphophonological alternation during training and whether the learning or generalization behavior is tested, I ran two experimental studies on learning vowel nasalization and on learning retroflexion by German native speakers, whose artificial language learning experiments were framed by typological surveys, a production experiment, perception experiments and an acoustic analysis. As vowel nasalization and retroflexion do not occur phonemically in German and as the nasalization of vowels of different height as well as the retroflexion of different consonants is of equal structural complexity, these phonological patterns are the perfect test cases for studying possible effects of phonetic details in phonological learning while omitting the confounding factors of previously criticized studies. In both studies, I investigated the learnability of three different patterns (nasalization of high, mid and low vowels, retroflexion of plosives, fricatives and lateral approximants). I tested the predictions for the learning and generalization behavior of various phonetic learning biases in two different tests (perception and production) and under two different training conditions (with and without evidence for a morphophonological alternation). The strength with which each phonetic factor affects the learning and generalization behavior, hence, gives insights into the interaction and the individual contribution of each phonetic detail in phonological learning.

My studies showed that while the typological distribution of a phonological pattern is not directly affected by phonetic biases, synchronic phonological learning is. Whereas the learning behavior is mainly affected by ease of perception – and by perceptual similarity to L1 if available –, the generalization behavior is mainly affected by phonetic similarity – and by perceptual similarity to L1 if available, or by ease of perception if no evidence for an alternation is provided during training. The mapping of a non-native sound onto a native sound category due to the perceptual similarity to L1 means a learning advantage but a generalization disadvantage because misperception leads to mislearning that hampers generalizations. An effect of ease of articulation could not be found in the results – independent of type of test. Due to a possible misperception there might also be evidence for a simplicity bias in the experiments on retroflexion.

These results contradict the assumptions concerning markedness based on phonetically based phonology (Archangeli & Pulleyblank, 1994; Hayes & Steriade, 2004; Wilson, 2006). They also speak against a bias based on ease of articulation. The learning results can best be explained by a learning bias based on ease of perception which is stronger than a bias based on ease of articulation and which even shows up in a production task. A division in a phonetic bias based on ease of perception and a weaker bias based on ease of articulation is thus necessary (Glewwe, 2019). The results further speak against a strong influence of ease of perception in generalizations. The generalization results of the participants can best be explained by a bias based on phonetic similarity, which is stronger than a bias based on ease of perception. Hence, the learning behavior and the generalization behavior are biased by different phonetic factors, which is best captured by distances of contrasts: The larger the perceptual distance between two sounds is, the better the contrast between these two sounds is learned. In generalizations, however, the learned contrast is best generalized to the contrast that has the smallest distance to the learned one.

Additionally, the results display that a phonetic bias is stronger when no evidence for a morphophonological alternation is provided during training, which means that the training is unreliable concerning the relation between presented forms. With evidence for a morphophonological alternation – meaning an increased reliability of a phonological pattern –, the phonetic bias becomes weaker, whereas L1-specific information are used by the learners to master the task of phonological learning. However, a closer inspection of how else the reliability of a training pattern can be captured and how this reliability affects how a phonological pattern is learned and generalized is still necessary. While there are already approaches which assume an effect of phonological similarity to already existing lexical representations, such as Stamer & Vitevitch (2012) and Storkel et al. (2006), these approaches do not fully fit to the findings of this dissertation as there are no existing lexical representations of the presented pseudowords in the artificial language learning experiments. My findings suggest a similarity effect independent of the lexical representations and further underline the importance of related forms and of allophones during phonological learning (Mitterer & Reinisch, 2017; Mitterer et al., 2013, 2018; Reinisch et al., 2014). Further research on the reliability of a pattern, the interaction between L1-specific and language-general perception and its contribution is necessary to model how the learning and the generalization behavior are affected by these factors.

Besides, this dissertation demonstrates general effects of phonetic details on phonological representations, which is important for the interface of phonetics and phonology. It was shown that phonetic details, such as phonetic similarity or ease of perception, affect how learners perform to trained and untrained items. Theories in which phonetics and phonology are separate modules cannot account for these effects. The results, thus, give evidence for the controversial role of phonetics in phonology and call for a tight integration of phonetics in phonology with phonological representations including phonetic details.

In conclusion, the present dissertation shows that phonological learning and phonological generalizations are affected by different phonetic factors as well as by L1-specific factors, if these are available. The dissertation gives experimental evidence for the existence and importance of phonetic biases based on perceptual ease, phonetic similarity and perceptual similarity to L1 in phonological learning. More research is necessary to model their effects on phonological learning on the one hand and on phonological generalizations on the other hand. In addition to that, the results of this dissertation call for a tight integration of phonetics in phonology. Theories must include phonetic details in phonological representations in such a way that they are able to account for the differences in learning and generalizing phonological patterns that differ in their phonetic details.

References

Abramson, A. & Lisker, L. (1970). Discriminability along the voicing continuum: Cross-language tests. In Hála, B., Romportl, M., & Janota, P. (Eds.), *Proceedings of the 6th International Congress of Phonetic Sciences*, (pp. 569–573)., Prague. Academia.

Ahmed, R. & Agrawal, S. (1969). Significant features in the perception of (Hindi) consonants. *Journal of the Acoustical Society of America*, *45*(3), 758–763.

Albright, A. & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, *90*(2), 119–161.

Aoyama, K., Flege, J. E., Guion, S. G., Akahane-Yamada, R., & Yamada, T. (2004). Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics*, *32*(2), 233–250.

Archangeli, D. & Pulleyblank, D. (1994). *Grounded phonology*. Cambridge, MA: MIT Press.

Azevedo, M. M. (2005). *Portuguese: A linguistic introduction*. Cambridge, UK: Cambridge University Press.

Baayen, R. H. (2008). *Analysing linguistic data (Vol. 505)*. Cambridge, UK: Cambridge University Press.

Baayen, R. H. (2013). *languageR: Data sets and functions with "Analyzing linguistic data: A practical introduction to statistics"*. R package version 1.4.1.

Baer-Henney, D. (2015). *Learner's little helper: Strength and weakness of the substantive bias in phonological acquisition*. PhD thesis, University of Potsdam.

Baer-Henney, D., Kügler, F., & van de Vijver, R. (2015a). The interaction of language-specific and universal factors during the acquisition of morphophonemic alternations with exceptions. *Cognitive Science*, *39*(7), 1537–1569.

Baer-Henney, D., Kügler, F., & van de Vijver, R. (2015b). The nature of the substantive bias. In Grillo, E. & Jepson, K. (Eds.), *Proceedings of the 39th annual Boston University Conference on Language Development*, volume 1, (pp. 62–74)., Somerville, MA. Cascadilla Press.

Baer-Henney, D. & van de Vijver, R. (2012). On the role of substance, locality, and amount of exposure in the acquisition of morphophonemic alternations. *Laboratory Phonology*, *3*(2), 221–249.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.1-8.

Becker, M. & Levine, J. (2010). Experigen – an online experiment platform. Available at https://github.com/tlozoot/experigen.

Becker, M., Nevins, A., & Levine, J. (2012). Asymmetries in generalizing alternations to and from initial syllables. *Language*, *88*(2), 231–268.

Beddor, P. S. (1993). The perception of nasal vowels. In M. K. Huffman & R. A. Krakow (Eds.), *Nasals, nasalization and the velum. Phonetics and phonology, volume 5* (pp. 171–196). San Diego, CA: Academic Press.

Beddor, P. S., Coetzee, A. W., Styler, W., McGowan, K. B., & Boland, J. E. (2018). The time course of individuals' perception of coarticulatory information is linked to their production: Implications for sound change. *Language*, *94*(4), 931–968.

Beddor, P. S., Krakow, R. A., & Goldstein, L. M. (1986). Perceptual constraints and phonological change: A study of nasal vowel height. *Phonology Yearbook*, *3*, 197–217.

Bell-Berti, F. (1993). Understanding velic motor control: Studies of segmental context. In M. K. Huffman & R. A. Krakow (Eds.), *Nasals, nasalization, and the velum. Phonetics and phonology, volume 5* (pp. 63–85). San Diego, CA: Academic Press.

Berent, I. (2013). The phonological mind. *Trends in Cognitive Sciences*, *17*(7), 319–327.

Berent, I., Everett, D. L., & Shimron, J. (2001). Do phonological representations specify variables? Evidence from the Obligatory Contour Principle. *Cognitive Psychology*, *42*(1), 1–60.

Berent, I., Marcus, G. F., Shimron, J., & Gafos, A. I. (2002). The scope of linguistic generalizations: Evidence from Hebrew word formation. *Cognition*, *83*, 113–139.

Berent, I., Steriade, D., Lennertz, T., & Vaknin, V. (2007). What we know about what we have never heard: Evidence from perceptual illusions. *Cognition*, *104*(3), 591–630.

Berent, I., Vaknin, V., & Marcus, G. F. (2007). Roots, stems, and the universality of lexical representations: Evidence from Hebrew. *Cognition*, *104*, 254–286.

Berent, I., Wilson, C., Marcus, G. F., & Bemis, D. (2012). On the role of variables in phonology: Remarks on Hayes and Wilson. *Linguistic Inquiry*, *43*(1), 97–119.

Berko Gleason, J. (1958). The child's learning of English morphology. *WORD*, *14*(2), 150–177.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). Timonium, MD: York Press.

Best, C. T. & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics*, *20*(3), 305–330.

Blevins, J. (2004). *Evolutionary phonology: The emergence of sound patterns*. Cambridge, UK: Cambridge University Press.

Blevins, J. (2017). Areal sound patterns: From perceptual magnets to stone soup. In R. Hickey (Ed.), *The Cambridge handbook of areal linguistics* (pp. 88–121). Cambridge, UK: Cambridge University Press.

Blevins, J. P. & Blevins, J. (2009). *Analogy in grammar: Form and acquisition*. Oxford, UK: Oxford University Press on Demand.

Boersma, P. & Weenink, D. (2019). Praat: Doing phonetics by computer (version 6.0.36) [computer program]. http://www.praat.org/.

Bond, Z. S. (1976). Identification of vowels excerpted from neutral and nasal contexts. *Journal of the Acoustical Society of America*, *59*(5), 1229–1232.

Booij, G. (1995). *The phonology of Dutch*. Oxford, UK: Clarendon Press.

Booij, G. (2010). *Construction morphology*. Oxford, UK: Oxford University Press.

Borg, A. & Azzopardi-Alexander, M. (1997). *Maltese*. Abingdon, UK: Routledge.

Bradlow, A., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, *101*(4), 2299–2310.

Broe, M. B. (1993). *Specification theory: The treatment of redundancy in generative phonology*. PhD thesis, University of Edinburgh.

Carignan, C. (2018). Using ultrasound and nasalance to separate oral and nasal contributions to formant frequencies of nasalized vowels. *Journal of the Acoustical Society of America*, *143*(5), 2588–2601.

Carignan, C., Shosted, R., Shih, C., & Rong, P. (2011). Compensatory articulation in American English nasalized vowels. *Journal of Phonetics*, *39*(4), 668–682.

Carpenter, A. C. (2006). *Acquisition of a natural versus an unnatural stress system*. Amherst, MA: University of Massachusetts Amherst.

Cebrian, J. (2006). Experience and the use of non-native duration in L2 vowel categorization. *Journal of Phonetics*, *34*(3), 372–387.

Chen, M. Y. (1995). Acoustic parameters of nasalized vowels in hearing-impaired and normal hearing speakers. *Journal of the Acoustical Society of America*, *98*(5), 2443–2453.

Chen, M. Y. (1997). Acoustic correlates of English and French nasalized vowels. *Journal of the Acoustical Society of America*, *102*(4), 2360–2370.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. & Halle, M. (1968). *The sound pattern of English*. New York City, NY: Harper and Row.

Clements, G. N. (1985). The geometry of phonological features. *Phonology Yearbook*, *2*(1), 225–252.

Clements, G. N. (1991). Place of articulation in consonants and vowels: A unified theory. *Working papers of the Cornell phonetics laboratory*, *5*, 77–123.

Clements, G. N. (2009). The role of features in phonological inventories. In E. Raimy & C. E. Cairns (Eds.), *Contemporary views on architecture and representations in phonological theory* (pp. 19–68). Cambridge, MA: MIT Press.

Clements, G. N. & Hume, E. V. (1995). The internal organization of speech sounds. In J. Goldsmith, J. Riggle, & A. C. L. Yu (Eds.), *The handbook of phonological theory* (pp. 245–306). Malden, MA: Blackwell.

Cristià, A., Mielke, J., Daland, R., & Peperkamp, S. (2013). Similarity in the generalization of implicitly learned sound patterns. *Laboratory Phonology*, *4*(2).

Cristià, A. & Seidl, A. (2008). Is infants' learning of sound patterns constrained by phonological features? *Language Learning and Development*, *4*(3), 203–227.

Croft, W. (2007). Construction grammar. In D. Geeraerts & H. Cuyckens (Eds.), *The Oxford handbook of cognitive linguistics* (pp. 463–508). Oxford, UK: Oxford University Press.

Crystal, D. (2008). *A dictionary of linguistics and phonetics, 6th edition*. Malden, MA: Blackwell Publishing.

Culbertson, J. (2012). Typological universals as reflections of biased learning: Evidence from artificial language learning. *Language and Linguistics Compass*, *6*(5), 310–329.

Daland, R., Hayes, B., Garellek, M., White, J., Davis, A., & Norrmann, I. (2011). Explaining sonority projection effects. *Phonology*, *28*(2), 197–234.

Dart, S. N. (1991). Articulatory and acoustic properties of apical and laminal articulations. *UCLA Working Papers in Phonetics*, *79*, 1–155.

Dart, S. N. & Nihalani, P. (1999). The articulation of Malayalam coronal stops and nasals. *Journal of the International Phonetic Association*, *29*(2), 129–142.

Delattre, P. (1954). Les attributs acoustiques de la nasalité vocalique et consonantique. *Studia Linguistica*, *8*(2), 103–109.

Delvaux, V. (2009). Perception du contraste de nasalité vocalique en français. *Journal of French Language Studies*, *19*(1), 25–59.

Demuth, K. (1995). Markedness and the development of prosodic structure. In J. Beckman (Ed.), *Proceedings of the North East Linguistic Society 25* (pp. 13–25). Amherst, MA: University of Massachusetts, Graduate Linguistics Student Association.

Dixon, R. M. W. (2011). *The languages of Australia*. Cambridge, UK: Cambridge University Press.

Draxler, C. (2014). Online experiments with the Percy software framework – experiences and some early results. In Calzolari, N., Choukri, K., Declerck, T., Loftsson, H., Maegaard, B., Mariani, J., Moreno, A., Odijk, J., & Piperidis, S. (Eds.), *Proceedings of the 9th International Conference on Language Resources and Evaluation*, (pp. 235–240)., Reykjavik. European Language Resources Association.

Earle, F. S., Landi, N., & Myers, E. B. (2017). Sleep duration predicts behavioral and neural differences in adult speech sound learning. *Neuroscience Letters, 636*, 77–82.

Eisenberg, P. & Sayatz, U. (2004). Left of number: Animacy and plurality in German nouns. In G. Müller, L. Gunkel, & G. Zifonun (Eds.), *Explorations in nominal inflection* (pp. 97–120). Berlin: Mouton de Gruyter.

Eisner, F. & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception & Psychophysics, 67*(2), 224–238.

Ernestus, M. & Baayen, R. H. (2003). Predicting the unpredictable: Interpreting neutralized segments in Dutch. *Language*, *79*(1), 5–38.

Escudero, P., Benders, T., & Lipski, S. C. (2009). Native, non-native and L2 perceptual cue weighting for Dutch vowels: The case of Dutch, German, and Spanish listeners. *Journal of Phonetics*, *37*(4), 452–465.

Fagyal, Z., Kibbee, D., & Jenkins, F. (2006). *French: A linguistic introduction*. Cambridge, UK: Cambridge University Press.

Fant, G. (1960). *The acoustic theory of speech production*. Den Haag: Mouton.

Finley, S. (2009). *Formal and cognitive restrictions on vowel harmony*. PhD thesis, Johns Hopkins University, Baltimore, MD.

Finley, S. (2012). Typological asymmetries in round vowel harmony: Support from artificial grammar learning. *Language and Cognitive Processes, 27*(10), 1550–1562.

Finley, S. (2017). Perceptually based constraints and metathesis: Evidence from artificial grammar. In Gallagher, G., Gouskova, M., & Yin, S. H. (Eds.), *Proceedings of the 2016 Annual Meetings on Phonology*, Washington, D.C. Linguistic Society of America.

Finley, S. & Badecker, W. (2007). Towards a substantively biased theory of learning. In Crane, T., David, O., Fenton, D., Haynie, H. J., Katseff, S., Lee-Goldman, R., Rouvier, R., & Yu, D. (Eds.), *Proceedings of 33rd Berkeley Linguistics Society*, (pp. 142–153)., Berkley, CA. Berkeley Linguistics Society.

Finley, S. & Badecker, W. (2009). Artificial language learning and feature-based generalization. *Journal of Memory and Language*, *61*(3), 423–437.

Finley, S. & Badecker, W. (2012). Learning biases for vowel height harmony. *Cognitive Science, 13*(3), 287–327.

Flege, J. E. (1991). Perception and production: The relevance of phonetic input to L2 phonological learning. In T. Hübner & C. A. Ferguson (Eds.), *Cross currents in second language acquisition and linguistic theory* (pp. 249–290). Amsterdam: John Benjamins.

Flege, J. E. (1995). Second language learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). Timonium, MD: York Press.

Flege, J. E., Bohn, O.-S., & Jang, S. (1997). Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics*, *25*(4), 437–470.

Flege, J. E., Takagi, N., & Mann, V. (1996). Lexical familiarity and English-language experience affect Japanese adults' perception of /ɹ/ and /l/. *Journal of the Acoustical Society of America*, *99*(2), 1161–1173.

Flemming, E. (2001). Scalar and categorical phenomena in a unified model of phonetics and phonology. *Phonology*, *18*(1), 7–44.

Fletcher, S. G., Sooudi, I., & Frost, S. D. (1974). Quantitative and graphic analysis of prosthetic treatment for "nasalance" in speech. *The Journal of Prosthetic Dentistry*, *32*(3), 284–291.

Glewwe, E. (2019). *Bias in phonotactic learning: Experimental studies of phonotactic implicationals*. PhD thesis, University of California, Los Angeles.

Gnanadesikan, A. E. (1993). The feature geometry of coronal subplaces. *University of Massachusetts Occasional Papers in Linguistics*, *16*(1), 27–67.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, *105*(2), 251–279.

Goldinger, S. D. & Azuma, T. (2004). Episodic memory reflected in printed word naming. *Psychonomic Bulletin & Review*, *11*(4), 716–722.

Gómez, D. M., Berent, I., Benavides-Varela, S., Bion, R. A. H., Cattarossi, L., Nespor, M., & Mehler, J. (2014). Language universals at birth. *Proceedings of the National Academy of Sciences*, *11*(6), 5837–5841.

Greenberg, J. H. (1950). The patterning of root morphemes in Semitic. *WORD*, *6*(2), 162–181.

Greenwood, A. (2016). *An experimental investigation of phonetic naturalness*. PhD thesis, University of California, Santa Cruz.

Greisbach, R. (2003). The pronunciation of loanwords in German. In Solé, M. J., Recasens, D., & Romero, J. (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, (pp. 799–802)., Barcelona.

Guion, S. G. (1996). *Velar palatalization: Coarticulation, perception, and sound change*. PhD thesis, University of Texas at Austin.

Guion, S. G. (1998). The role of perception in the sound change of velar palatalization. *Phonetica*, *55*(1-2), 18–52.

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, *107*(5), 2711–2724.

Hajek, J. (1997). *Universals of sound change in nasalization*. Malden, MA: Blackwell Publishing.

Hajek, J. & Maeda, S. (2000). Investigating universals of sound change: The effect of vowel height and duration on the development of distinctive nasalization. In M. B. Broe & J. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V* (pp. 52–69). Cambridge, UK: Cambridge University Press.

Hale, M. & Reiss, C. (2000). "Substance abuse" and "dysfunctionalism": Current trends in phonology. *Linguistic Inquiry*, *31*(1), 157–169.

Hallé, P. A., Best, C. T., & Levitt, A. (1999). Phonetic vs. phonological influences on French listeners: Perception of American English approximants. *Journal of Phonetics*, *27*(3), 281–306.

Hamann, S. (2003a). Norwegian retroflexion – licensing by cue or prosody? In Dahl, A., Bentzen, K., & Svenonius, P. (Eds.), *Special Issue of the Proceedings of the 19th Scandinavian Conference of Linguistics*, (pp. 63–77)., Nordlyd.

Hamann, S. (2003b). *The phonetics and phonology of retroflexes*. PhD thesis, University of Utrecht.

Hamp, E. P. (1996). On the Indo-European origins of the retroflexes in Sanskrit. *Journal of the American Oriental Society*, *116*(4), 719–724.

Hayes, B. (1999). Phonetically-driven phonology: The role of optimality theory and inductive grounding. In M. Darnell, E. A. Moravcsik, M. Noonan, F. J. Newmeyer, & K. Wheatley (Eds.), *Functionalism and formalism in Linguistics: Volume 1* (pp. 243–285). Amsterdam: John Benjamins.

Hayes, B. (2009). *Introductory phonology*. Malden, MA: Wiley-Blackwell.

Hayes, B. & Steriade, D. (2004). Introduction: The phonetic bases of phonological markedness. In B. Hayes, D. Steriade, & R. M. Kirchner (Eds.), *Phonetically based phonology* (pp. 1–33). Cambridge, UK: Cambridge University Press.

Hayes, B. & Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, *39*(3), 379–440.

Hecker, M. H. L. (1962). Studies of nasal consonants with an articulatory speech synthesizer. *Journal of the Acoustical Society of America*, *34*(2), 179–188.

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., & Kliegl, R. (2011). dlexdb – eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau*, *62*(1), 10–20.

Hintzman, D. L. (1984). Minerva 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*, *16*(2), 96–101.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95*(4), 528–551.

Hope, R. M. (2013). *Rmisc: Ryan Miscellaneous*. R package version 1.5.

House, A. S. (1957). Analog studies of nasal consonants. *Journal of Speech and Hearing Disorders, 22*(2), 190–204.

House, A. S. & Stevens, K. N. (1956). Analog studies of the nasalization of vowels. *Journal of Speech and Hearing Disorders, 21*(2), 218–232.

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

Jackendoff, R. & Audring, J. (2016). Morphological schemas. *The Mental Lexicon, 11*(3), 467–493.

Jun, J. (2004). Place assimilation. In B. Hayes, D. Steriade, & R. M. Kirchner (Eds.), *Phonetically based phonology* (pp. 58–86). Cambridge, UK: Cambridge University Press.

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language, 33*(5), 630–645.

Jusczyk, P. W., Smolensky, P., & Allocco, T. (2002). How English-learning infants respond to markedness and faithfulness constraints. *Language Acquisition, 10*(1), 31–73.

Kapatsinski, V. (2013a). Conspiring to mean: Experimental and computational evidence for a usage-based harmonic approach to morphophonology. *Language, 89*(1), 110–148.

Kapatsinski, V. (2013b). Morphological schema induction by means of conditional inference trees. In Cartoni, B., Bernhard, D., & Tribout, D. (Eds.), *TACMO workshop. Theoretical and* *computational morphology: New trends and synergies*, (pp. 11–14)., Geneva.

Kaun, A. (2004). The phonetic foundations of the rounding harmony typology. In B. Hayes, D. Steriade, & R. M. Kirchner (Eds.), *Phonetically based phonology* (pp. 87–116). Cambridge, UK: Cambridge University Press.

Keating, P. A. (1984). Phonetic and phonological representation of stop consonant voicing. *Language*, *60*(2), 286–319.

Keating, P. A. (1991). Coronal places of articulation. In C. Paradis & J.-F. Prunet (Eds.), *The special status of coronals: Internal and external evidence* (pp. 29–48). Amsterdam: Elsevier.

Keating, P. A., Garellek, M., & Kreiman, J. (2015). Acoustic properties of different kinds of creaky voice. In The Scottish Consortium for ICPhS 2015 (Ed.), *Proceedings of the 18th International Congress of Phonetic Sciences*, (pp. paper 821)., Glasgow, UK. University of Glasgow.

Keating, P. A. & Lahiri, A. (1993). Fronted velars, palatalized velars, and palatals. *Phonetica*, *50*(2), 73–101.

Kimper, W. (2016). Asymmetric generalisation of harmony triggers. In Hansson, G. O., Farris-Trimble, K., McMullin, K., & Pulleyblank, D. (Eds.), *Proceedings of the 2015 Annual Meeting on Phonology*, Washington, D.C. Linguistic Society of America.

Kingston, J. (2007). The phonetics-phonology interface. In P. de Lacy (Ed.), *The Cambridge handbook of phonology* (pp. 401–434). Cambridge, UK: Cambridge University Press.

Kingston, J. & Macmillan, N. A. (1995). Integrality of nasalization and F1 in vowels in isolation and before oral and nasal consonants: A detection-theoretic application of the Garner paradigm. *Journal of the Acoustical Society of America*, *97*(2), 1261–1285.

Knoblauch, K. (2014). *psyphy: Functions for analyzing psychophysical data in R*. R package version 0.1-9.

Kohler, K. J. (1990). Segmental reduction in connected speech in German: Phonological facts and phonetic explanations. In W. J. Hardcastle & A. Marchal (Eds.), *Speech production and speech modelling* (pp. 69–92). New York City, NY: Springer.

Komaki, R., Akahane-Yamada, R., & Choi, Y. (1999). Effects of native language on the perception of American English /r/ and /l/: Cross-language comparison between Korean and Japanese. *Technical report of The Institute of Electronics, Information and Communication Engineers*, (SP 99-45), 39–46.

Köpcke, K.-M. (1988). Schemas in German plural formation. *Lingua*, *74*(4), 303–335.

Kopkallı, H. (1993). *A phonetic and phonological analysis of final devoicing in Turkish.* PhD thesis, University of Michigan.

Kraljic, T. & Samuel, A. G. (2005). Perceptual learning for speech: Is there a return to normal? *Cognitive Psychology*, *51*(2), 141–178.

Kraljic, T. & Samuel, A. G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review*, *13*(2), 262–268.

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

Krull, D. (1990). Relating acoustic properties to perceptual responses: A study of Swedish voiced stops. *Journal of the Acoustical Society of America*, *88*(6), 2557–2570.

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*(2), 93–107.

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

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2018). *lmerTest: Tests in linear mixed effects models*. R package version 3.1-1.

Ladefoged, P. (2001). *A course in phonetics (4th ed.)*. Boston, MA: Heinle & Heinle.

Ladefoged, P. & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford, UK: Blackwell.

Ladefoged, P. & Wu, Z. (1984). Places of articulation: An investigation of Pekingese fricatives and affricates. *Journal of Phonetics, 12*(3), 267–278.

Laeufer, C. (2010). Nasal vowels in French loanwords in German: The effect of linguistic environment. *Folia Linguistica, 44*(1), 53–101.

Laver, J. (1994). *Principles of phonetics*. Cambridge, UK: Cambridge University Press.

Lehiste, I. (1970). *Suprasegmentals*. Cambridge, MA: MIT Press.

Lehiste, I. (1988). *Lectures on language contact*. Cambridge, MA: MIT Press.

Lindblad, P. & Lundqvist, S. (1999). How and why do the tongue gestures of [t],[d],[l],[n],[s], and [r] differ. In Ohala, J. J., Hasegawa, Y., Ohala, M., Granville, D., & Bailey, A. C. (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences,* (pp. 417–420)., San Francisco, CA.

Linzen, T. & Gallagher, G. (2017). Rapid generalization in phonotactic learning. *Laboratory Phonology*, *8*(1), 24.

Luce, P. A. & Pisoni, D. B. (1998). Recognizing spoken words: The neighborhood activation model. *Ear and Hearing, 19*(1), 1–36.

MacKay, I. R. A., Flege, J. E., Piske, T., & Schirru, C. (2001). Category restructuring during second-language speech acquisition. *Journal of the Acoustical Society of America*, *110*(1), 516–528.

Macmillan, N. A., Kingston, J., Thorburn, R., Walsh Dickley, L., & Bartels, C. (1999). Integrality of nasalization and F1. II. Basic sensitivity and phonetic labeling measure distant sensory and decision-rule interactions. *Journal of the Acoustical Society of America*, *106*(5), 2913–2932.

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

Marcus, G. F. (2001). *The algebraic mind: Integrating connectionism and cognitive science*. Cambridge, MA: MIT Press.

Martin, A. & Peperkamp, S. (2020). Phonetically natural rules benefit from a learning bias: A re-examination of vowel harmony and disharmony. *Phonology*, *37*(1), 65–90.

Martin, A. & White, J. (2021). Vowel harmony and disharmony are not equivalent in learning. *Linguistic Inquiry*, *52*(1), 227–239.

McCarthy, J. J. (1988). Feature geometry and dependency: A review. *Phonetica*, *45*(2-4), 84–108.

McCloy, D. R. (2016). *phonR: Tools for phoneticians and phonologists*. R package version 1.0-7.

Merkx, M., Rastle, K., & Davis, M. H. (2011). The acquisition of morphological knowledge investigated through artificial language learning. *Quarterly Journal of Experimental Psychology*, *64*(6), 1200–1220.

Mermelstein, P. (1977). On detecting nasals in continuous speech. *Journal of the Acoustical Society of America*, *61*(2), 581–587.

Mintz, T. H. (2002). Category induction from distributional cues in an artificial language. *Memory & Cognition*, *30*(5), 678–686.

Mitterer, H. & Reinisch, E. (2017). Surface forms trump underlying representations in functional generalisations in speech perception: The case of German devoiced stops. *Language, Cognition and Neuroscience*, *32*(9), 1133–1147.

Mitterer, H., Reinisch, E., & McQueen, J. (2018). Allophones, not phonemes in spoken-word recognition. *Journal of Memory and Language*, *98*, 77–92.

Mitterer, H., Sharenborg, O., & McQueen, J. M. (2013). Phonological abstraction without phonemes in speech perception. *Cognition*, *129*, 356–361.

Moran, S. & McCloy, D. R. (2019). Phoible 2.0. Jena: Max Planck Institute for the Science of Human History. https://phoible.org/.

Moreton, E. (2008). Analytic bias and phonological typology. *Phonology*, *25*(1), 83–127.

Moreton, E. (2012). Inter- and intra-dimensional dependencies in implicit phonotactic learning. *Journal of Memory and Language*, *67*(1), 165–183.

Moreton, E. & Pater, J. (2012a). Structure and substance in artificial phonology learning, part I: Structure. *Language and Linguistics Compass*, *6*(11), 686–701.

Moreton, E. & Pater, J. (2012b). Structure and substance in artificial phonology learning, part II: Substance. *Language and Linguistics Compass*, *6*(11), 702–718.

Morgan, J. L. & Newport, E. L. (1981). The role of constituent structure in the induction of an artificial language. *Journal of Verbal Learning and Verbal Behavior*, *20*(1), 67–85.

Narayanan, S. & Kaun, A. (1999). Acoustic modeling of Tamil retroflex liquids. In Ohala, J. J., Hasegawa, Y., Ohala, M., Granville, D., & Bailey, A. C. (Eds.), *Proceedings of the 14th International Congress of Phonetic Sciences,* (pp. 2097–2101)., San Francisco, CA.

Nielsen, K. (2011). Specificity and abstractness of VOT imitation. *Journal of Phonetics, 39*(2), 132–142.

Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47,* 204–238.

Ohala, J. J. (1975). Phonetic explanations for nasal sound patterns. In C. A. Ferguson, L. M. Hyman, & J. J. Ohala (Eds.), *Nasálfest: Papers from a symposium on nasals and nasalization* (pp. 289–316). Standford, CA: Stanford University, Linguistics Department.

Ohala, J. J. (1986). Consumer's guide to evidence in phonology. *Phonology Yearbook, 3,* 3–26.

Ohala, J. J. (1993). Sound change as nature's speech perception experiment. *Speech Communication, 13*(1), 155–161.

Ohala, M. & Ohala, J. J. (2001). Acoustic VC transitions correlate with degree of perceptual confusion of place contrast in Hindi. In N. Grø & J. Rischel (Eds.), *Travaux du Cercle Linguistique de Copenhague* (pp. 265–284). Copenhagen: Reitzel.

Öhman, S. (1966). Perception of segments of VCCV utterances. *Journal of the Acoustical Society of America, 40*(5), 979–988.

Pardo, J. S. (2006). On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America, 119*(4), 2382–2393.

Pardo, J. S., Jay, I. C., & Krauss, R. M. (2010). Conversational role influences speech imitation. *Attention, Perception, & Psychophysics, 72*(8), 2254–2264.

Pater, J. & Moreton, E. (2012). Structurally biased phonology: Complexity in learning and typology. *Journal of the English and Foreign Languages University, Hyderabad, 3*(2), 1–44.

Peirce, J. W. (2007). PsychoPy – Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1), 8–13.

Peperkamp, S. & Dupoux, E. (2007). Learning the mapping from surface to underlying representations in an artificial language. *Laboratory Phonology*, *9*, 315–338.

Peperkamp, S. & Martin, A. (2016). Sleep-dependent consolidation in the learning of natural vs. unnatural phonological rules. Presentation at the 15th Conference on Laboratory Phonology, Ithaca, NY.

Peperkamp, S., Skoruppa, K., & Dupoux, E. (2006). The role of phonetic naturalness in phonological rule acquisition. In Bamman, D., Magnitskaia, T., & Zaller, C. (Eds.), *Proceedings of the 30th annual Boston University Conference on Language Development*, volume 2, (pp. 464–475)., Somerville, MA. Cascadilla Press.

Polka, L. (1995). Linguistic influences in adult perception of non-native contrasts. *Journal of the Acoustical Society of America*, *97*(2), 1286–1296.

Pruthi, T. & Epsy-Wilson, C. (2004). Acoustic parameters for automatic detection of nasal manner. *Speech Communication, 43*(3), 225–239.

Pulleyblank, D. (2011). Vowel height. In M. v. Oostendorp, C. Ewen, E. V. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology* (pp. 491–518). Malden, MA: Blackwell.

Pulleyblank, E. G. (1989). The role of coronal in articulator based features. In *Proceedings of the 25th Conference of the Chicago Linguistic Society*, (pp. 379–393)., Chicago, IL. Chicago Linguistic Society.

Pycha, A., Nowak, P., Shin, E., & Shosted, R. (2003). Phonological rule-learning and its implications for a theory of vowel harmony. In Garding, G. & Tsujimura, M. (Eds.), *Proceedings of the 22nd West Coast Conference on Formal Linguistics0*, (pp. 101–114)., Somerville, MA. Cascadilla Press.

R Core Team (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org.

Rauber, A. S., Escudero, P., Bion, R. A. H., & Baptista, B. O. (2005). The interrelation between the perception and production of English vowels by native speakers of Brazilian Portuguese. In *Proceedings of the 9th European Conference on Speech Communication and Technology*, (pp. 2913–2916).

Reinisch, E., Wozny, D. R., Mitterer, H., & Holt, L. L. (2014). Phonetic category recalibration: What are the categories? *Journal of Phonetics*, *45*, 91–105.

Röttger, T., Winter, B., Grawunder, S., Kirby, J., & Grice, M. (2014). Assessing incomplete neutralization of final devoicing in German. *Journal of Phonetics*, *43*, 11–25.

Saffran, J. & Thiessen, E. (2003). Pattern induction by infant language learners. *Developmental Psychology*, *39*(3), 484–494.

Sancier, M. L. & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics*, *25*(4), 421–436.

Sarmā, D. (2004). *Tribal languages of Ladakh (Part III)*. New Delhi: Mittal Publications.

Schiel, F. (2010). BAStat: New statistical resources at the Bavarian Archive for Speech Signals. In Calzolari, N., Choukri, K., Maegaard, B., Mariani, J., Odijk, J., Piperidis, S., Rosner, M., & Tapias, D. (Eds.), *Proceedings of the 7th International Conference on Language Resources and Evaluation*, (pp. paper 277)., Valletta, Malta. European Language Resources Association.

Schwartz, M. F. (1968). The acoustics of normal and nasal vowel production. *Cleft Palate Journal*, *5*(2), 125–140.

Seidl, A. & Buckley, E. (2005). On the learning of arbitrary phonological rules. *Language Learning and Development*, *1*(3–4), 289–316.

Sheldon, A. & Strange, W. (1982). The acquisition of /r/ and /l/ by Japanese learners of English: Evidence that speech production can precede speech perception. *Applied Psycholinguistics*, *3*(3), 243–261.

Shosted, R., Carignan, C., & Rong, P. (2012). Managing the distinctiveness of phonemic nasal vowels: Articulatory evidence from Hindi. *Journal of the Acoustical Society of America*, *131*(1), 455–465.

Skoruppa, K., Lambrechts, A., & Peperkamp, S. (2011). The role of phonetic distance in the acquisition of phonological alternations. In Lima, S., Mullin, K., & Smith, B. (Eds.), *Proceedings of the 39th North Eastern Linguistics Conference*, (pp. 717–729)., Somerville, MA. Cascadilla Press.

Skoruppa, K. & Peperkamp, S. (2011). Adaptation to novel accents: Feature-based learning of context-sensitive phonological regularities. *Cognitive Science*, *35*(2), 348–366.

Slowiaczek, L. M. & Dinnsen, D. A. (1985). On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics*, *13*(3), 325–341.

Snodgrass, J. G. & Vanderwart, M. (1980). A standardised set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, *6*(2), 174–215.

Stamer, M. K. & Vitevitch, M. S. (2012). Phonological similarity influences word learning in adults learning Spanish as a foreign language. *Bilingualism*, *15*(3), 490–502.

Stausland Johnsen, S. (2012). From perception to phonology: The emergence of perceptually motivated constraint rankings. *Lingua*, *122*(2), 125–143.

Steriade, D. (1997). Phonetics in phonology: The case of laryngeal neutralization. Ms., University of California at Los Angeles.

Steriade, D. (2001a). Directional asymmetries in place assimilation: A perceptual account. In E. V. Hume & K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 219–250). San Diego, CA: Academic Press.

Steriade, D. (2001b). The phonology of perceptibility effects: The P-map and its consequences for constraint organization. Ms., University of California at Los Angeles.

Steriade, D. (2009). The phonology of perceptibility effects: The P-map and its consequences for constraint organization. In K. Hanson & S. Inkelas (Eds.), *The nature of the word* (pp. 150–178). Cambridge, MA: MIT Press.

Stevens, K. N. (1998). *Acoustic phonetics*. Cambridge, MA: MIT Press.

Stevens, K. N. & Blumstein, S. E. (1975). Quantal aspects of consonant production and perception: A study of retroflex stop consonants. *Journal of Phonetics*, *3*(4), 215–233.

Stevens, K. N., Keyser, S. J., & Kawasaki, H. (1986). Toward a phonetic and phonological theory of redundant features. In J. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 426–449). Hillsdale, NJ: Lawrence Erlbaum Associates.

Storkel, H. L., Armbrüster, J., & Hogan, T. P. (2006). Differentiating phonotactic probability and neighborhood density in adult word learning. *Journal of Speech, Language, and Hearing Research*, *49*(6), 1175–1192.

Styler, W. (2015). *On the acoustical and perceptual features of vowel nasality*. PhD thesis, University of Colorado at Boulder.

Styler, W. (2017). On the acoustical features of vowel nasality in English and French. *Journal of the Acoustical Society of America*, *142*(4), 2469–2482.

Styler, W. & Scarborough, R. (2017). Nasality automeasure script package. Boulder, CO: University of Colorado's Phonetics Lab. Available at https://github.com/stylerw/styler_praat_scripts.

Takagi, N. (1993). *Perception of American English /r/ and /l/ by adult Japanese learners of English: A unified view*. PhD thesis, University of California, Irvine.

Tang, K. & Baer-Henney, D. (under review). Disentangling L1 and L2 effects in artificial language learning. Preprint: https://dx.doi.org/10.31234/osf.io/df8ru.

Trask, R. (1996). *A dictionary of phonetics and phonology*. Abingdon, UK: Routledge.

Uffmann, C. (2011). The organization of features. In M. v. Oostendorp, C. Ewen, E. V. Hume, & K. Rice (Eds.), *The Blackwell companion to phonology* (pp. 1–26). Malden, MA: Blackwell.

van de Vijver, R. & Baer-Henney, D. (2012). Sonority intuitions are provided by the lexicon. In S. Parker (Ed.), *The sonority controversy*, number 18 in Phonology & Phonetics (pp. 193–213). Berlin: Mouton de Gruyter.

van de Vijver, R. & Baer-Henney, D. (2014). Developing biases. *Frontiers in Psychology*, *5*(634).

Vance, T. (1987). *An introduction to Japanese phonology*. Albany, NY: State University of New York Press.

Wängler, H.-H. (1964). *Atlas deutscher Sprachlaute*. Berlin: Akademie-Verlag.

Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, *32*(2), 251–276.

Westbury, J. R. & Keating, P. A. (1986). On the naturalness of stop consonant voicing. *Journal of Linguistics*, *22*(1), 145–166.

White, J. (2013). *Bias in phonological learning: Evidence from saltation*. PhD thesis, University of California, Los Angeles.

White, J. (2014). Evidence for a learning bias against saltatory phonological alternations. *Cognition*, *130*(1), 96–115.

White, J. (2017). Accounting for the learnability of saltation in phonological theory: A maximum entropy model with a P-map bias. *Language*, *1*(93), 1–36.

White, J. & Sundara, M. (2014). Biased generalization of newly learned phonological alternations by 12-month-old infants. *Cognition*, *133*(1), 85–90.

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. New York City, NY: Springer.

Wickham, H., François, R., Henry, L., & Müller, K. (2019). *dplyr: A grammar of data manipulation*. R package version 0.8.1.

Wiese, R. (1996). *The phonology of German*. Oxford, UK: Oxford University Press.

Wilson, C. (2003). Experimental investigation of phonological naturalness. In Garding, G. & Tsujimura, M. (Eds.), *Proceedings of the 22nd West Coast Conference on Formal Linguistics 22*, (pp. 533–546)., Somerville, MA. Cascadilla Press.

Wilson, C. (2006). Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science*, *30*(5), 945–982.

Wright, J. T. (1975). Effects of vowel nasalization on the perception of vowel height. In C. A. Ferguson, L. M. Hyman, & J. J. Ohala (Eds.), *Nasálfest: Papers from a symposium on nasals and nasalization* (pp. 373–388). Stanford, CA: Stanford University, Linguistics Department.

Wurzel, W. U. (1965). *Flexionsmorphologie und Natürlichkeit*. Berlin: Akademie-Verlag.

Yu, A. C. L. (2004). Explaining final obstruent voicing in Lezgian: Phonetics and history. *Language, 80*(1), 73–97.

Zellou, G., Dahan, D., & Embick, D. (2017). Imitation of coarticulatory vowel nasality across words and time. *Language, Cognition and Neuroscience*, *32*(6), 776–791.

Zellou, G., Scarborough, R., & Nielsen, K. (2016). Phonetic imitation of coarticulatory vowel nasalization. *Journal of the Acoustical Society of America*, *140*(5), 3560–3575.

Zhang, H. (2015). Production and acoustics of creaky nasal vowels. Master's thesis, University of Colorado at Boulder.

Zhang, J. (2004). The role of constrast-specific and language-specific phonetics in contour tone distribution. In B. Hayes, D. Steriade, & R. M. Kirchner (Eds.), *Phonetically based phonology* (pp. 157–190). Cambridge, UK: Cambridge University Press.

Zsiga, E. C. (2013). *The sounds of language: An introduction to phonetics and phonology*. Malden, MA: Wiley-Blackwell.

Zuraw, K. (2007). The role of phonetic knowledge in phonological patterning: Corpus and survey evidence from Tagalog infixation. *Language*, *83*(2), 277–316.

Appendix

Table 26: Stimuli for the production experiment (see chapter 2.2): IPA transcription and German orthography.

IPA transcription			German orthography		
CV [m] V	**CVCV**	**CVCV [m]**	**CV [m] V**	**CVCV**	**CVCV [m]**
duma	duba	dubam	duma	duba	dubam
dumɛ	dubɛ	dubɛm	dumä	dubä	dubäm
dumi	dubi	dubim	dumi	dubi	dubim
dumo	dubo	dubom	dumo	dubo	dubom
dumu	dubu	dubum	dumu	dubu	dubum
kuma	kufa	kufam	kuma	kufa	kufam
kumɛ	kufɛ	kufɛm	kumä	kufä	kufäm
kumi	kufi	kufim	kumi	kufi	kufim
kumo	kufo	kufom	kumo	kufo	kufom
kumu	kufu	kufum	kumu	kufu	kufum
poma	poga	pogam	poma	poga	pogam
pomɛ	pogɛ	pogɛm	pomä	pogä	pogäm
pomi	pogi	pogim	pomi	pogi	pogim
pomo	pogo	pogom	pomo	pogo	pogom
pomu	pogu	pogum	pomu	pogu	pogum
ʃoma	ʃota	ʃotam	schoma	schota	schotam
ʃomɛ	ʃotɛ	ʃotɛm	schomä	schotä	schotäm
ʃomi	ʃoti	ʃotim	schomi	schoti	schotim
ʃomo	ʃoto	ʃotom	schomo	schoto	schotom
ʃomu	ʃotu	ʃotum	schomu	schotu	schotum
vuma	vusa	vusam	wuma	wusa	wusam
vumɛt	vusɛ	vusɛm	wumä	wusä	wusäm
vumi	vusi	vusim	wumi	wusi	wusim
vumo	vuso	vusom	wumo	wuso	wusom
vumu	vusu	vusum	wumu	wusu	wusum

Table 27: Stimuli (in IPA transcription) in the training phase of the artificial language learning experiment (see chapter 2.4) and in the exposure phase of the productive task testing vowel nasalization (see chapter 3): singular forms.

[a] -learners	**[ɛ] -learners**	**[i] -learners**
doba	dobɛ	dobi
poba	pobɛ	pobi
ʃuba	ʃubɛ	ʃubi
dofa	dofɛ	dofi
kufa	kufɛ	kufi
ʃofa	ʃofɛ	ʃofi
doga	dogɛ	dogi
kuga	kugɛ	kugi
voga	vogɛ	vogi
duta	dutɛ	duti
pota	potɛ	poti
ʃuta	ʃutɛ	ʃuti
doza	dozɛ	dozi
kuza	kuzɛ	kuzi
ʃoza	ʃozɛ	ʃozi
vuza	vuzɛ	vuzi

Table 28: Stimuli (in IPA transcription) in the training phase of the artificial language learning experiment (see chapter 2.4) and in the exposure phase of the productive task testing vowel nasalization (see chapter 3): plural forms.

[a] -learners	**[ɛ] -learners**	**[i] -learners**
dobãm	dobɛ̃m	dobĩm
pobãm	pobɛ̃m	pobĩm
ʃubãm	ʃubɛ̃m	ʃubĩm
dofãm	dofɛ̃m	dofĩm
kufãm	kufɛ̃m	kufĩm
ʃofãm	ʃofɛ̃m	ʃofĩm
dogãm	dogɛ̃m	dogĩm
kugãm	kugɛ̃m	kugĩm
vogãm	vogɛ̃m	vogĩm
dutãm	dutɛ̃m	dutĩm
potãm	potɛ̃m	potĩm
ʃutãm	ʃutɛ̃m	ʃutĩm
dozãm	dozɛ̃m	dozĩm
kuzãm	kuzɛ̃m	kuzĩm
ʃozãm	ʃozɛ̃m	ʃozĩm
vuzãm	vuzɛ̃m	vuzĩm

Table 29: Stimuli (in IPA transcription) in the training phase of the artificial language learning experiment (see chapter 2.4) and in the exposure phase of the productive task testing vowel nasalization (see chapter 3): diminutive forms.

[a] -learners	[ɛ] -learners	[i] -learners
dobal	dobɛl	dobil
pobal	pobɛl	pobil
ʃubal	ʃubɛl	ʃubil
dofal	dofɛl	dofil
kufal	kufɛl	kufil
ʃofal	ʃofɛl	ʃofil
dogal	dogɛl	dogil
kugal	kugɛl	kugil
vogal	vogɛl	vogil
dutal	dutɛl	dutil
potal	potɛl	potil
ʃutal	ʃutɛl	ʃutil
dozal	dozɛl	dozil
kuzal	kuzɛl	kuzil
ʃozal	ʃozɛl	ʃozil
vuzal	vuzɛl	vuzil

Table 30: Stimuli (in IPA transcription) in the test phase of the artificial language learning experiment testing vowel nasalization (see chapter 2.4): correct form ∼ incorrect form.

V_2 = [a]	**V_2 = [ɛ]**	**V_2 = [i]**
potãm ∼ potam	dobɛ̃m ∼ dobɛm	ʃofĩm ∼ ʃofim
dofãm ∼ dofam	dozɛ̃m ∼ dozɛm	dogĩm ∼ dogim
kugãm ∼ kugam	kufɛ̃m ∼ kufɛm	dutĩm ∼ dutim
ʃubãm ∼ ʃubam	ʃutɛ̃m ∼ ʃutɛm	kuzĩm ∼ kuzim
kogãm ∼ kogam	vofɛ̃m ∼ vofɛm	vobĩm ∼ vobim
ʃotãm ∼ ʃotam	ʃobɛ̃m ∼ ʃobɛm	dotĩm ∼ dotim
dubãm ∼ dubam	vubɛ̃m ∼ vubɛm	vufĩm∼ vufim
pufãm ∼ pufam	kutɛ̃m ∼ kutɛm	puzĩm ∼ puzim
dobal ∼ dobãl	pobɛl ∼ pobɛ̃l	dofil ∼ dofĩl
dogal ∼ dogãl	ʃozɛl ∼ ʃozɛ̃l	vogil ∼ vogĩl
kufal ∼ kufãl	dutɛl ∼ dutɛ̃l	ʃubil ∼ ʃubĩl
ʃutal ∼ ʃutãl	kuzɛl ∼ kuzɛ̃l	kugil ∼ kugĩl
vozal ∼ vozãl	kobɛl ∼ kobɛ̃l	ʃogil ∼ ʃogĩl
kotal ∼ kotãl	pofɛl ∼ pofɛ̃l	kozil ∼ kozĩl
pubal ∼ pubãl	dugɛl ∼ dugɛ̃l	dufil ∼ dufĩl
vutal ∼ vutãl	ʃuzɛl ∼ ʃuzɛ̃l	putil ∼ putĩl

Dies ist eine Geschichte über die beiden Freundinnen Sabine und Claudia.
Letzte Woche hat sich Sabine verletzt, als sie versuchte, auf einem Baumstamm zu **balancieren**.
Sie ist dabei mit ihrem Fuß umgeknickt.
Ihre beste Freundin Claudia hat sie daraufhin zum Arzt begleitet und für diese Hilfe möchte sich Sabine **revanchieren**.
Sie überlegt lange hin und her, bis ihr letztlich eine gute Idee kommt.
Als Dankeschön hat Sabine ein Geschenk für Claudia: ihr liebstes **Parfum**.
Darüber wird sich Claudia bestimmt freuen.
Außerdem möchte Sabine für Claudia kochen und hat deshalb ein Treffen mit ihr **arrangiert**.
Am Samstagabend kommt Claudia Sabine in ihrer neuen Wohnung besuchen.
Sabine hat Claudias Lieblingsspeise gekocht, und zwar **Kartoffel-Gratin**.
Claudia freut sich sehr darüber.
Zum Nachtisch gibt es Schokomuffins mit einer Dekoration aus **Fondant**.
Es sind kleine Blumen mit Zuckerperlen.
Claudia liebt Schokolade und **Bonbons**.
Auch Sabine kann Süßigkeiten kaum widerstehen.
Während des Essens bewundert Claudia Sabines **Interieur**.
Sie hat ihre neue Wohnung wirklich schön dekoriert.
Nach dem Essen setzen sich die beiden auf den **Balkon**.
Dieser ist zwar recht klein, aber trotzdem groß genug für die beiden Freundinnen.
Dort erzählt Claudia von Thomas, ihrem **Cousin**.
Thomas hat früher als Lehrer gearbeitet.
Aber er ist seit kurzem **Pensionär**.
Nun hat er viel Freizeit.
Er möchte sich gerne ehrenamtlich **engagieren**.
So kann er seine freie Zeit sinnvoll nutzen.
Allerdings weiß er noch nicht genau, in welcher **Branche**.
Claudia fragt, ob Sabine vielleicht eine Idee hätte.
Sabine erzählt daraufhin von dem leeren **Waggon**.
Den hat sie letzten Monat erst entdeckt.
Er steht in der Nähe des Bahnhofs und ist der Sitz des gemeinnützigen Vereins 'Jedem eine **Chance**!'.

Figure 40: Sentences for the reading task in the productive task testing vowel nasalization (see chapter 3). French loan words are printed in bold for illustrative purposes here. In the experiment they were not highlighted.

Table 31: Stimuli (in IPA transcription) in the training phase of the artificial language learning experiment testing retroflexion (see chapter 4.4): singular forms.

[l] -learners	[s] -learners	[t] -learners
dɛl	dɛs	dɛt
dɔl	dɔs	dɔt
fɛl	fɛs	fɛt
fɪl	fɪs	fɪt
fɔl	fɔs	fɔt
gɪl	gɪs	gɪt
gɔl	gɔs	gɔt
kɛl	kɛs	kɛt
kɪl	kɪs	kɪt
pɛl	pɛs	pɛt
pɪl	pɪs	pɪt
pɔl	pɔs	pɔt
bɪl	bɪs	bɪt
bʊl	bʊs	bʊt
vɛl	vɛs	vɛt
vʊl	vʊs	vʊt

Table 32: Stimuli (in IPA transcription) in the training phase of the two artificial language learning experiments testing retroflexion (see chapters 4.4 and 5): plural forms.

[l] -learners	**[s] -learners**	**[t] -learners**
dɛɭɾa	dɛʂɾa	dɛʈɾa
dɔɭɾa	dɔʂɾa	dɔʈɾa
fɛɭɾa	fɛʂɾa	fɛʈɾa
fɪɭɾa	fɪʂɾa	fɪʈɾa
fɔɭɾa	fɔʂɾa	fɔʈɾa
gɪɭɾa	gɪʂɾa	gɪʈɾa
gɔɭɾa	gɔʂɾa	gɔʈɾa
kɛɭɾa	kɛʂɾa	kɛʈɾa
kɪɭɾa	kɪʂɾa	kɪʈɾa
pɛɭɾa	pɛʂɾa	pɛʈɾa
pɪɭɾa	pɪʂɾa	pɪʈɾa
pɔɭɾa	pɔʂɾa	pɔʈɾa
bɪɭɾa	bɪʂɾa	bɪʈɾa
bʊɭɾa	bʊʂɾa	bʊʈɾa
vɛɭɾa	vɛʂɾa	vɛʈɾa
vʊɭɾa	vʊʂɾa	vʊʈɾa

Table 33: Stimuli (in IPA transcription) in the training phase of the two artificial language learning experiments testing retroflexion (see chapters 4.4 and 5): diminutive forms.

[l] -learners	[s] -learners	[t] -learners
dɛlna	dɛsna	dɛtna
dɔlna	dɔsna	dɔtna
fɛlna	fɛsna	fɛtna
fɪlna	fɪsna	fɪtna
fɔlna	fɔsna	fɔtna
gɪlna	gɪsna	gɪtna
gɔlna	gɔsna	gɔtna
kɛlna	kɛsna	kɛtna
kɪlna	kɪsna	kɪtna
pɛlna	pɛsna	pɛtna
pɪlna	pɪsna	pɪtna
pɔlna	pɔsna	pɔtna
bɪlna	bɪsna	bɪtna
bʊlna	bʊsna	bʊtna
vɛlna	vɛsna	vɛtna
vʊlna	vʊsna	vʊtna

Table 34: Stimuli (in IPA transcription) in the test phase of the artificial language learning experiment testing retroflexion (see chapter 4.4): correct form ∼ incorrect form.

C_2 = [l]	C_2 = [s]	C_2 = [t]
pɛɭɾa ∼ pɛlɾa	dɛʂɾa ∼ dɛsɾa	vɛʈɾa ∼ vɛtɾa
pɪɭɾa ∼ pɪlɾa	fɔʂɾa ∼ fɔsɾa	kɪʈɾa ∼ kɪtɾa
bɛɭɾa ∼ bɛlɾa	kɔʂɾa ∼ kɔsɾa	gɛʈɾa ∼ gɛtɾa
vɔɭɾa ∼ vɔlɾa	kʊʂɾa ∼ kʊsɾa	fʊʈɾa ∼ fʊtɾa
bʊɭɾa ∼ bʊlɾa	vʊʂɾa ∼ vʊsɾa	pɔʈɾa ∼ pɔtɾa
gɔɭɾa ∼ gɔlɾa	gɪʂɾa ∼ gɪsɾa	bʊʈɾa ∼ bʊtɾa
vɪɭɾa ∼ vɪIɾa	bɛʂɾa ∼ bɛsɾa	dɪʈɾa ∼ dɪtɾa
pʊɭɾa ∼ pʊlɾa	dɪʂɾa ∼ dɪsɾa	kɔʈɾa ∼ kɔtɾa
pɛlna ∼ pɛɭna	dɛsna ∼ dɛʂna	vɛtna ∼ vɛʈna
fɪlna ∼ fɪɭna	fɔsna ∼ fɔʂna	kɪtna ∼ kɪʈna
bɛlna ∼ bɛɭna	bɛsna ∼ bɛʂna	gɛtna ∼ gɛʈna
vɔlna ∼ vɔɭna	kʊsna ∼ kʊʂna	fʊtna ∼ fʊʈna
bʊlna ∼ bʊɭna	vʊsna ∼ vʊʂna	pɔtna ∼ pɔʈna
gɔlna ∼ gɔɭna	gɪsna ∼ gɪʂna	bʊtna ∼ bʊʈna
vɪlna ∼ vɪɭna	kɔsna ∼ kɔʂna	dɪtna ∼ dɪʈna
pʊlna ∼ pʊɭna	dɪsna ∼ dɪʂna	kɔtna ∼ kɔʈna

Table 35: Stimuli (in IPA transcription) in the test phase of the artificial language learning experiment testing retroflexion without alternation (see chapter 5): correct form ∼ incorrect form.

C_2 = [l]	C_2 = [s]	C_2 = [t]
pɛɭɾa ∼ pɛlɾa	dɛʂɾa ∼ dɛsɾa	vɛʈɾa ∼ vɛtɾa
pɪɭɾa ∼ pɪlɾa	fɔʂɾa ∼ fɔsɾa	kɪʈɾa ∼ kɪtɾa
bɛɭɾa ∼ bɛlɾa	kɔʂɾa ∼ kɔsɾa	gɛʈɾa ∼ gɛtɾa
vɔɭɾa ∼ vɔlɾa	kʊʂɾa ∼ kʊsɾa	fʊʈɾa ∼ fʊtɾa
bʊɭɾa ∼ bʊlɾa	vʊʂɾa ∼ vʊsɾa	pɔʈɾa ∼ pɔtɾa
gɔɭɾa ∼ gɔlɾa	gɪʂɾa ∼ gɪsɾa	bʊʈɾa ∼ bʊtɾa
vɪɭɾa ∼ vɪIɾa	bɛʂɾa ∼ bɛsɾa	dɪʈɾa ∼ dɪtɾa
pʊɭɾa ∼ pʊlɾa	dɪʂɾa ∼ dɪsɾa	kɔʈɾa ∼ kɔtɾa

Lebenslauf

Kristina Kim Strütjen

Beruflicher Werdegang

seit 11/2015	Wissenschaftliche Mitarbeiterin Institut für Sprache und Information Heinrich-Heine-Universität Düsseldorf
04/2018 – 09/2018	Lehrbeauftragte Institut für Linguistik Universität Stuttgart
10/2015	Studentische Hilfskraft Institut für Sprache und Information Heinrich-Heine-Universität Düsseldorf
10/2014 – 07/2015	Studentische Hilfskraft Institut für Sprache und Information Heinrich-Heine-Universität Düsseldorf

Schulische und akademische Ausbildung

10/2015 – 09/2021	Heinrich-Heine-Universität Düsseldorf Promotion im Fach Allgemeine Sprachwissenschaft
10/2013 – 09/2015	Heinrich-Heine-Universität Düsseldorf Studiengang: Linguistik Abschluss: Master of Arts
10/2010 – 09/2013	Heinrich-Heine-Universität Düsseldorf Studiengang: Linguistik integrativ Abschluss: Bachelor of Arts
2001 – 2010	Vestisches Gymnasium Kirchhellen, Bottrop Abschluss: Allgemeine Hochschulreife

Eidesstattliche Erklärung

Ich versichere an Eides Statt, dass die Dissertation von mir selbständig und ohne unzulässige fremde Hilfe unter Beachtung der "Ordnung über die Grundsätze zur Sicherung guter wissenschaftlicher Praxis an der Heinrich-Heine-Universität Düsseldorf" erstellt worden ist.

Düsseldorf, 15.11.2021

(Ort, Datum)

(Unterschrift)