# Phonological Contrast and Conflict in Dutch Vowels:

## Neurobiological and Psycholinguistic Evidence from Children and Adults

**Nadine de Rue**

The research reported in this dissertation was supported by the Netherlands Research Council (NWO), project number 322-75-005.

This project has also received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (Grant agreement number: 695481).

ISBN: 978-94-6421-574-8
Cover and layout: Bregje Jaspers | proefschriftontwerp.nl
Print: Ipskamp Printing, Enschede

# Phonological Contrast and Conflict in Dutch Vowels:

## Neurobiological and Psycholinguistic Evidence from Children and Adults

Proefschrift ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. J.H.J.M. van Krieken,
volgens besluit van het college voor promoties
in het openbaar te verdedigen op

donderdag 27 januari 2022
om 10.30 uur precies

door

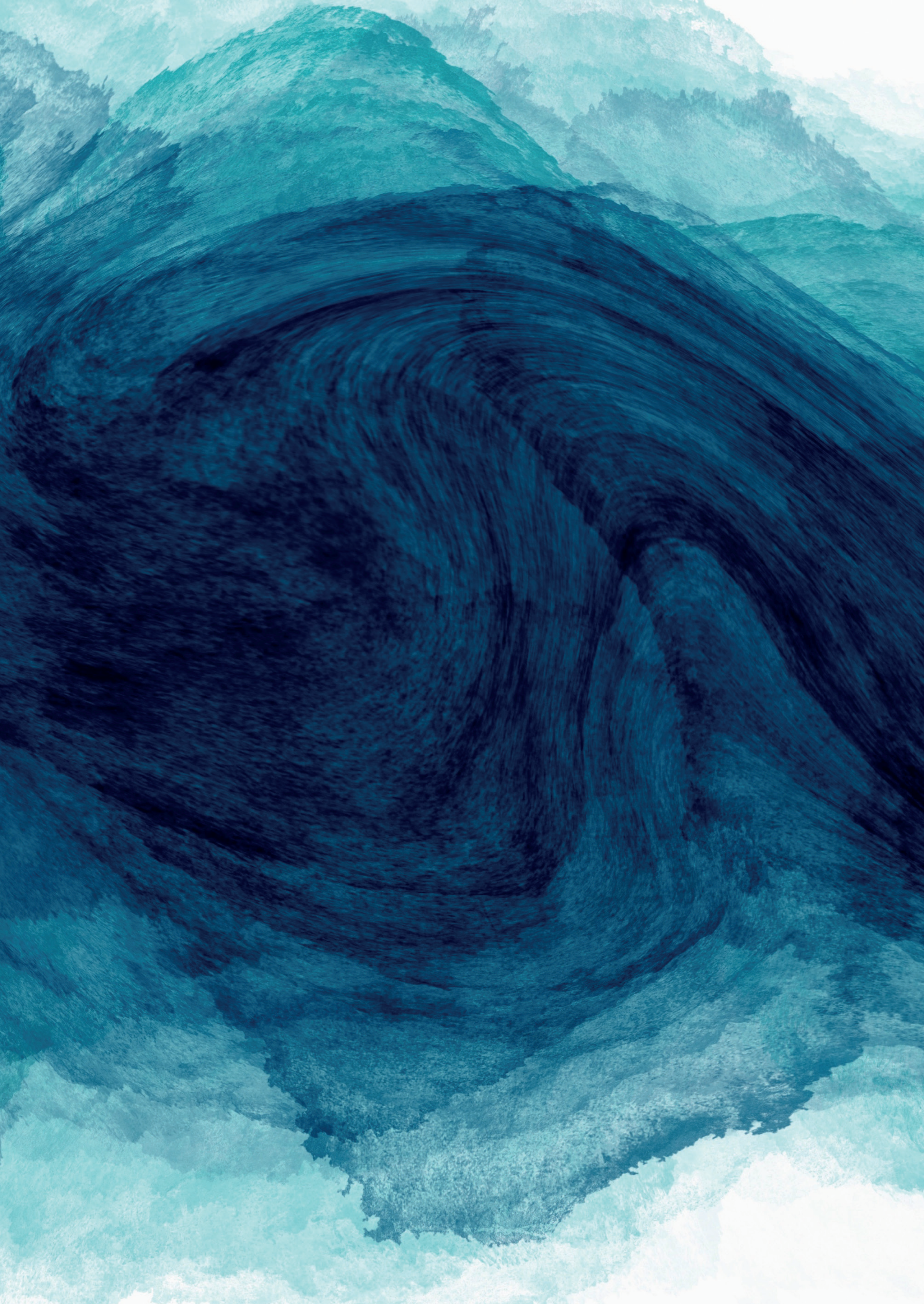
**Nadine Paula Willem Daniëlle de Rue**
geboren op 12 juli 1990
te Venlo

**Promotoren:**
Prof. dr. J.P.M. Fikkert
Prof. dr. A. Lahiri (University of Oxford, Verenigd Koninkrijk)

**Copromotor:**
Dr. T.M. Snijders (MPI)

**Manuscriptcommissie:**
Prof. dr. M. van Oostendorp
Prof. dr. H.M.G.M. Jacobs
Prof. dr. J. Grijzenhout (Universiteit Leiden)
Prof. dr. M. Scharinger (Philipps-Universität Marburg, Duitsland)
Dr. S.M. Brouwer

## CONTENTS



# Chapter 1

## General introduction

In language comprehension, the ability to recognize words is essential. The Dutch words *boer* (/bur/; 'farmer') and *buur* (/byr/; 'neighbour') only differ with respect to a single speech sound (/u/ vs. /y/), but have distinct meanings. The distinction between /u/ and /y/ is only a matter of tongue position: /u/ is *back*, whereas /y/ is *front*. Sensitivity to such *phonological contrasts* is vital for successful word recognition. However, there is a substantial amount of acoustic variation in normal speech: no word is ever pronounced alike twice. Thus, not all acoustic differences are meaningful. If acoustic variation does not signal a different word, it can be ignored for word recognition. Thus, what complicates word recognition is that some differences in acoustics are important to notice, whilst others can be ignored.

Each phoneme (or speech sound) in a language has a unique combination of phonological features. A phoneme can be *front* or *back* as in the example mentioned above. In the case of vowels this means that the tongue constriction is in the front or back of the oral cavity. A vowel can also be *high* or *low*. This height refers to a high or low tongue position in the oral cavity. Each phoneme differs from all other phonemes with respect to at least one feature. Sensitivity to these featural differences enables us to detect meaningful differences during speech recognition.

Differences between vowels are primarily due to differences in tongue and/or lip position. Lips may be round like /o/ in *boot* (/bot/; 'boat') or not, like /e/ in *been* (/ben/; 'leg'). The tongue can be high, as in /i/ in *fiets* (/fits/; 'bike'), or low, like /a/ in *laars* (/lars/; 'boot'). The tongue can also be front (e.g. /y/ like *buur*) or back (e.g. /u/ like *boer*). The space within which the tongue can move in height (high-low) and place (front-back) is referred to as the vowel space. Vowel space can be seen as a stylized representation of the oral cavity. The stylized representation of vowel space is given in Figure 1.1.

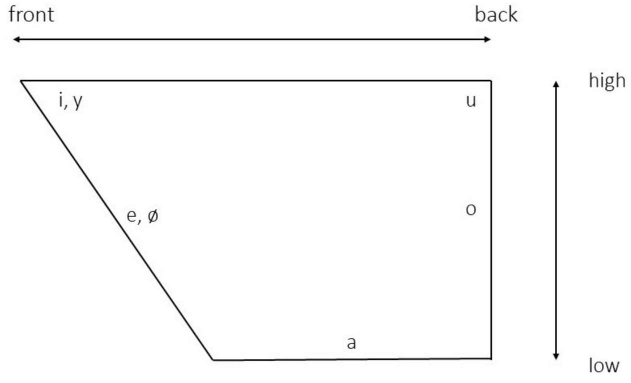



**Figure 1.1** Stylized vowel space.

In order to recognize words, a listener has to compare the acoustic input (i.e. speech) to the words stored in one's mind (i.e. the representations of words in the mental lexicon). A key issue in phonology is what information these representations contain, and how they enable us to efficiently recognize words. In other words: how are phonological contrasts perceived and non-distinctive differences ignored?

Some phonological features are universal, which means that they occur in all languages. Other contrasts are language specific, and only occur in some languages. For example, the contrast between front and back vowels is language universal (99.56% of languages in the UPSID database; Maddieson, 1984): all languages have contrastive front and back vowels like /i/ and /u/ respectively. In many languages, the features front and back coincide with lip rounding: back vowels are typically round, and front vowels are not. In Dutch, however, frontness and roundedness do not coincide. In addition to the language universal contrast between front and back vowels, Dutch also has a language specific roundedness contrast in front vowels, e.g. front /i/ like in *bier* (/bir/; 'beer') versus front round /y/ like in *buur* (/byr/; 'neighbour'). Dutch is one of the roughly 7% of languages that have this contrast (Maddieson, 1984).

As a result, Dutch has a relatively rare three-way vowel contrast between front /e/ (e.g. *steen:* /sten/; 'rock'), front round /ø/ (e.g. *neus:* /nøs/; 'nose'), and back round /o/ (e.g. *boot:* /bot/; 'boat'). Similarly, German also has a contrast between /e/, /ø/ and /o/. Representations of these vowels in German have received attention in the literature (e.g. Wiese, 1996; Eulitz & Lahiri, 2004; Cornell, Lahiri & Eulitz, 2011; Scharinger, Lahiri & Eulitz, 2010). The similarity in vowel inventory suggests that the same underlying representations may be at stake in both languages. However, this is not a given. Vowels may sound the same, but have different feature representations (or vice versa), as they configurate differently from language to language (Dresher, 2009). An example included in Dresher (2009) is the following: Czech, Slovak, and Russian vowels appear largely similar on the surface, while the shape of the inventory as well as sound patterns imply that different phonological contrasts are made, as becomes clear in Jakobson (1931/62). Even though the /a/ of Slovak and the /a/ of Czech are phonetically nearly identical, Jakobson considers it to pattern as a back vowel in Slovak, because there are two low vowels: /a/ and /ä/, the latter being more front (or more 'acute' in Jakobson's terms). In Czech, Jakobson assumes only height to be contrastive for /a/, since there are no other low vowels in Czech. As the front/back dimension appears contrastive in Slovak based on the two low vowels /a/ and /ä/, non-low vowels like /i/ and /u/ may be distinguished based on this front/back dimension also, and lip rounding is regarded not distinctive. Furthermore, Jakobson (1931/62) assumes that due to the existence of front/back allophones in non-low Russian vowels, only one dimension (lip rounding) is contrastive, whereas for Czech he suggests the two dimensions front/back and lip rounding (or in Jakobsons' terms 'acute/grave' and 'flatness') are important to distinguish non-low vowels. As such, vowel inventories and sound patterning play a major role in which features are distinctive.

Indeed, there are reasons to believe similar representations of front/back and roundedness in German and Dutch are not evident when looking at the linguistic systems as a whole. First, in German, many phonetically front round vowels arise because of morphological umlaut, which fronts back vowels in certain plurals, diminutives, and adjectival and verbal forms. For example, German has the singular-plural alternation *V/o/gel – V[ø]gel* ('bird – birds'), whereas Dutch has *vogel – vogels*, without vowel alternation. Although the classic interpretation of Germanic umlaut sketched by Twaddell (1938) regards umlaut as a fundamentally phonological process, umlaut is no longer phonologically transparent in modern German, and has become a morphological rule (Wiese, 1996). The /ø/ is therefore often morphologically derived, and might not be part of the underlying representation of stems that may alternate between front round and back round vowels (Scharinger, 2009). Modern Dutch does not have morphological umlaut, and front round vowels are truly part of the Dutch vowel inventory. Furthermore, the roundedness contrast in vowels may be more important in Dutch than in German. Non-alternating stems with front round vowels are much more frequent in Dutch, and Dutch words with front round vowels more often form minimal pairs with other words with respect to lip rounding. As such, roundedness may play a different role in German and Dutch. This may be reflected in differences in phonological representation and processing.

In the current dissertation, I focus on how the contrasts between front/back and round/unround are represented, enabling Dutch listeners to perceive these phonological contrasts in vowels. I will refer to the front/back contrast as the *place contrast* and to the contrast between round and unround as the *labiality contrast.* I will use experimental evidence from speech perception to gain insight into what information phonological representations contain, and how they are employed, in order to handle the place and labiality vowel contrast in Dutch (for example, a three-way distinction between front /e/, front round /ø/, and back round /o/).

## The Dutch vowel system

Standard Dutch as spoken in the Netherlands differs in pronunciation from the variety spoken in Belgium (e.g. Adank, Van Hout & Smits, 2004). This dissertation investigates phonological representations of Standard Dutch as spoken in the Netherlands, henceforth called Dutch.

Dutch has twelve monophthongs /ɑ, ɛ, ɪ, ɔ, ʏ, a, e, i, o, ø, u, y/, three diphthongs /ɛi, œy, ɑu/ (like in *wijn;* 'wine', *buik;* 'belly', and *oud;* 'old' respectively), and a neutral or reduced vowel /ə/ (schwa) (Booij, 1995). Since the literature does not provide consensus regarding phonological features of the Dutch vowels, Figure 1.2 depicts the monophthongs in vowel space based on average formant frequencies reported by Adank et al. (2004). Formants are distinctive frequency components of the acoustic signal, related to the vocal tract shape (e.g. tongue and lip position).

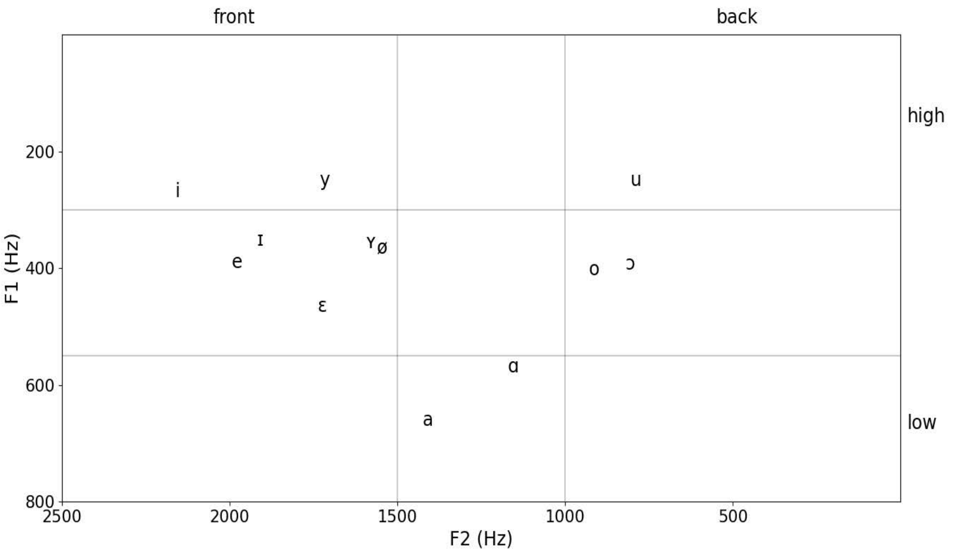



**Figure 1.2** Dutch vowel space. Twelve Dutch monophthongs are plotted in vowel space based on average first and second formant frequencies (F1 and F2) reported in Hertz (Hz) for male speakers of Northern Standard Dutch in Adank et al. (2004). High F1 frequency relates to low position of the tongue. High F2 frequency is related to front position of the tongue. Compare to Figure 1.1.

Two categories of monophthongs are distinguished, referred to in various ways, such as tense vs. lax, or long vs. short vowels, or ATR vs. RTR (advanced tongue root vs. retracted tongue root). For example, the /e/ in *peen* ('carrot') is categorized as tense (or long) and /ɛ/ in *pen* ('pen') is categorized as lax (or short). I do not intend to engage in the discussion about what the proper terminology should be. Rather, following Van Oostendorp (1995) who argues that the tense/lax terminology is a more suitable description for Dutch vowels than long vs. short, I will use the term *tense* to refer to the group of vowels /a, e, i, o, ø, u, y/, and the term *lax* to refer to the vowels /ɑ, ɛ, ɪ, ɔ, ʏ/.

Tense vowels /e/, /o/ and /ø/ may be considered phonetically semi-diphthongized (Adank et al., 2004) in Dutch since they are typically produced with a slight formant shift, but phonologically they are considered monophthongs as diphthongization is not obligatory.

Dutch has three front round vowels, namely /ʏ/ like in *put* ('well'), /y/ like in *vuur* ('fire'), and /ø/ like in *neus* ('nose'). They occur rather frequently and form minimal pairs with front vowels like /i/ and /e/ as well as back round vowels like /u/ and /o/. Hence, front round vowels should be distinguished from front unround vowels and back vowels. How this is done is of central interest in this dissertation. Below, I discuss the relevant phonological features for the contrasts of interest: lip rounding (henceforth: *labiality contrast*), and front/back (henceforth: *place contrast*).

Although focus lies on tense vowels /e/, /ø/ and /o/, the three lax vowels /ɛ/ (like in *pet;* 'cap'), /ʏ/ (like in *put;* 'well'), and /ɔ/ (like in *pot;* 'jar') are also included in the experiment in Chapter 3. In particular with respect to /ʏ/, we cannot avoid discussing height features as well. First, labiality and place are discussed simultaneously, as they are not always regarded separate contrasts in the literature. Subsequently, height is discussed separately.

**Labiality and place**

Different approaches have been suggested to deal with the three-way contrast concerning labiality and place (for example /e/, /ø/ and /o/). Zwaardemaker and Eijkman (1928) and De Groot (1931) consider a single phonological feature *backness* with three degrees (front, mid and back), see Table 1.1. Lip rounding is not taken into account as a separate feature.

**Table 1.1** Vowel chart following Zwaardemaker and Eijkman (1928) with some irrelevant omissions.

| | **front** | **mid** | **back** |
|---|---|---|---|
| closed | y i | | u |
| half-closed | e | ø | o |
| half-open | e | ʏ | ɔ |
| open | | a | |

Note that in Table 1.1 /y/ and /i/ are not distinguished based on backness and height. This issue is solved in the system proposed by De Groot (1931), who considers the three vowels /y/, /ø/ and /ʏ/ to be mid, so front vowels are consistently considered nonlabial: /i/, /e/ and /ɛ/. /u/, /o/ and /ɔ/ are back vowels (see Figure 1.3 in the next section about height).

Similar to De Groot (1931), Moulton (1962) also splits the front, front round and back vowels into three categories. The crucial difference compared to De Groot is that Moulton uses two dimensions rather than one to define the categories. Where De Groot used three levels of a single dimension called backness, Moulton uses the two dimensions (1) front – back, and (2) spread – rounded (see the vowel chart in Table 1.2 below) to make a three-way distinction. Both dimensions (or contrasts) are binary: there are two mutually exclusive counterparts assumed as features. Both dimensions are used separately, i.e. they do not coincide, indicating both distinctions are important. Although *spread* only occurs with front, *rounded* may be combined with either front or back. In a way, Moulton thus splits place (front-back) and labiality (*rounded* and *spread*) into two separate contrasts.

**Table 1.2** Simplified version of the vowel chart by Moulton (1962).

| | **front spread** | **front rounded** | **back rounded** |
|---|---|---|---|
| high | i | y | u |
| higher-mid | e | ø | o |
| lower-mid | e | ʏ | ɔ |
| low | | a, ɑ | |

In his generative phonology of Dutch, Booij (1981) also assumes bivalent features – following the Sound Pattern of English (SPE; Chomsky & Halle, 1968). Rather than Moulton's *spread* and *rounded*, however, Booij uses the feature [± ROUND] to indicate the difference between labial and nonlabial (i.e. round and unround) vowels. Thus, he considers the vowels /y, u, ø, ʏ, ɔ, o/ to be [+ROUND]. Also, similar to SPE, Booij (1981) uses the feature [±BACK] to split vowel space into front and back. Front vowels are [-BACK]. In Dutch, these are /y, i, ɪ, e, ɛ, ø, ʏ/. Back vowels are [+BACK]. According to Booij, these are Dutch vowels: /a, ɑ, o, u, ɔ/. Thus, accounting for labiality and place contrasts, Booij assumes [+ROUND] to combine with either [+BACK] or [-BACK].

Gussenhoven (1992) also provides a description of the Dutch vowels, albeit less detailed. Similar to Booij (1981), Gussenhoven (1992) assumes vowels to split into two categories regarding the front/back dimension. Labiality is considered a separate dimension. However, he assumes monovalent features: a segment either has a feature or it does not. No positive or negative counterparts are considered. As such, rather than [-ROUND] like in Booij (1981) (or SPE) or spread like in Moulton (1962), nonlabial is the absence of a feature.

**Height**

The literature shows controversy regarding the number of height levels assumed for the Dutch vowels as well as regarding the placement of the vowels on these levels. Particularly, the vowel /ʏ/ (in for example *put*; 'well') is subject to debate.

De Groot (1931) assumes three height levels, both for tense vowels and for lax vowels – see vowel charts in Figure 1.3 and Figure 1.4 respectively.

| | | |
|---|---|---|
| /i/ | /y/ | /u/ |
| /e/ | /ø/ | /o/ |
| | /a/ | |

**Figure 1.3** Tense vowels on three height levels by De Groot (1931). Height in vertical direction, backness in horizontal direction.

| | | |
|---|---|---|
| /ɪ/ | | /ɔ/ |
| | /ʏ/ | |
| /ɛ/ | | /ɑ/ |

**Figure 1.4** Lax vowels on three height levels by De Groot (1931). Height in vertical direction, backness in horizontal direction.

Placement of vowels in the vowel charts in Figure 1.3 and 1.4 implies /ʏ/ is the lax counterpart of /ø/. Tense vowels /e/, /ø/ and /o/ share their height level: they are all mid. /ʏ/ is also mid, but /ɛ/ is low and /ɔ/ is high.

Zwaardemaker and Eijkman (1928) and Eijkman (1937) consider four rather than three height levels. Relative placement of tense vowels /e/, /ø/ and /o/ is similar to De Groot (1931): all three are on the same height level, albeit this level is now called half-closed (see Table 1.1, already included in the previous section). In contrast to De Groot (1931), Zwaardemaker and Eijkman place /ʏ/ on the same height level as /ɛ/ and /ɔ/. However, Zonneveld and Trommelen (1980) point out a controversy regarding /ʏ/ in Zwaardemaker and Eijkman (1928) and Eijkman (1937). In Zwaardemaker and Eijkman (1928) /ʏ/ is described slightly more open than the vowel which occurs independently in a longer version in some loanwords such as *freule* (fr[ʏ:]le; 'lady'), whilst Eijkman (1937) describes it as slightly more closed than that vowel.

Similar to Zwaardemaker and Eijkman (1928) and Eijkman (1937), Moulton (1962) also assumes four height levels in Dutch – see above Table 1.2. In this classification, tense vowels /e/, /ø/ and /o/ are again placed on the same height level: all three are considered higher-mid. Although using a different terminology (i.e. lower-mid vs. half-open), Moulton assumes /ʏ/ is placed on the same height as /ɛ/ and /ɔ/, which is similar to Zwaardemaker and Eijkman (1928) and Eijkman (1937). Interestingly, Zwaardemaker and Eijkman (1928) as well as Moulton (1962) place lax vowels lower than tense vowels.

Booij (1981) also assumes four levels of height. He does so by combining two binary height features [±HIGH] and [±MID]. /e/, /ø/, and /o/ are considered high-mid vowels, defined as [+HIGH] and [+MID]. Lax vowels /ɛ/ and /ɔ/ are considered low-mid by the features [-HIGH] [+MID]. /ʏ/ is considered higher, namely on the same level as the tense vowels. This is referred to as high-mid in Booij (1981) and half-close in Booij (1995), which is essentially the same.

Booij (1981) states that he assumes four height levels for Dutch vowels, because others have argued that this is necessary in some languages other than Dutch. However, four levels are not actually required in Dutch to make all the required distinctions. A system with height features HIGH and LOW, in combination with distinguishing three levels of height, would be sufficient, as is for example assumed by Rietveld, Kerkhoff and Gussenhoven (2004). They assume three levels of height, with /ʏ/ considered as high. Tense vowels /e/, /ø/, and /o/ as well as lax vowels /ɛ/ and /ɔ/ lack [HIGH] and [LOW], and hence are mid vowels (neither high or low).

In sum, the literature includes much discussion about labiality, place and height features of Dutch vowels. Regarding **labiality** and **place**, there seems to be agreement about the fact that there is a three-way distinction between for instance /e/, /ø/ and /o/. Clearly, in Dutch, labiality is important in addition to place. In the literature, everyone makes this distinction, but they do so in different ways. Some assume a single dimension with three levels (e.g. De Groot, 1931), whereas others use two dimensions which separate place and labiality from each other (e.g. Moulton, 1962; Booij, 1981). This may be done using bivalent features like Booij (1981) does, or monovalent features such as suggested by Gussenhoven (1992). With respect to **height**, tense vowels /e/, /ø/ and /o/ are consistently assumed to all share the same height level, albeit in a four level system (e.g. Moulton, 1962; Booij, 1981) or a three level system (e.g. De Groot, 1931; Rietveld et al., 2004). They are always considered mid (meaning not high and not low); mid in a three level system, or close-mid (= high-mid) in a four level system. Height of the lax vowels /ɛ/, /ʏ/, and /ɔ/, however, remains controversial. Some assume these vowels to share their height level, whereas others assume they do not. Also, some place them on the same level as the tense vowels, whereas others do not. However, /ɛ/ and /ɔ/ are typically placed together on the same level. /ʏ/ may be considered similar, or higher, or lower than the other two lax vowels.

In this dissertation I investigate the underlying phonology of the Dutch vowel system with respect to the place contrast and the labiality contrast. Since there does not seem to be any controversy regarding height of tense vowels /e/, /ø/ and /o/, I consider these vowels a logical place to start to investigate how place and labiality are accounted for in Dutch. All experiments discussed in this dissertation hence focus on these three vowels.

In addition to the relevant phonological features, also their geometry and the nature of the representations are important to consider when investigating how vowels are distinguished in perception. Both geometry and representations are therefore discussed in the following two sections.

## Feature geometry

Different phonological theories make different assumptions regarding feature geometry. A feature geometry represents distinctive features as a structured hierarchy rather than as a matrix or set. With respect to the articulation of vowels (i.e. both place and labiality), three important types of hierarchies have been proposed in the literature regarding the features [LABIAL] (= round), [CORONAL] (= front) and [DORSAL] (= back):

1. [LABIAL], [CORONAL], and [DORSAL] are placed underneath a single place of articulation node;
2. The place of articulation node is split: [LABIAL] and [DORSAL] are grouped together, separated from [CORONAL].
3. The place of articulation node is split: [CORONAL] and [DORSAL] are grouped together, separated from [LABIAL].

I will argue that the latter seems to be the case for Dutch.

The first option - a single place of articulation node parenting [LABIAL], [CORONAL] and [DORSAL] - is assumed by, for instance, Clements and Hume (1995), similar to what is assumed by others (Chomsky & Halle, 1968; Sagey, 1986; Clements, 1992; Steriade, 1995). Clements and Hume (1995) assume the same features for vowels and consonants (i.e. unified features), although vowel features are placed on a separate level from consonantal features: the feature tree for consonants is duplicated to account for vowels, as can be seen below in Figure 1.5.

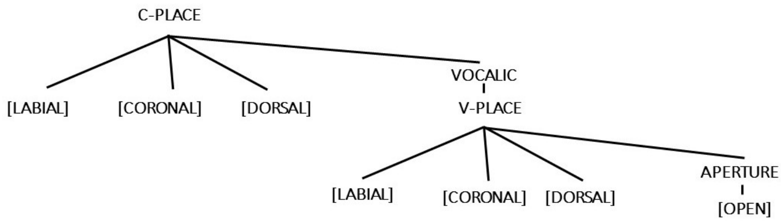



**Figure 1.5** Place of articulation feature geometry based on Clements & Hume (1995). C-PLACE refers to consonants and V-PLACE refers to vowels.

Another example of a geometry placing [LABIAL], [CORONAL] and [DORSAL] under the same parental node is the Featurally Underspecified Lexicon (FUL) model (e.g. Lahiri & Reetz, 2002, 2010; Kotzor, Wetterlin & Lahiri, 2017; Lahiri, 2018). Similar to Clements and Hume (1995), the same features are assumed for consonants and vowels, but in FUL these are unified within the same geometry; no duplication is assumed – both consonants and vowels use the same geometry. The geometry is given in Figure 1.6. Lahiri and Reetz (2002, 2010) assume an ARTICULATOR node with daughters [LABIAL], [CORONAL] and [DORSAL].

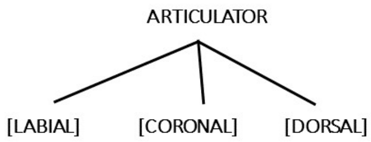



**Figure 1.6** Relevant features [LABIAL] , [CORONAL], and [DORSAL] in FUL's feature geometry.

In general, FUL assumes that features underneath the same parental node are mutually exclusive. This means that features below the same node may not be combined within one phoneme. Yet, as already noted by Lahiri (2018), this is not always the case for vowels. In Dutch, it is important to distinguish front and back vowels (i.e. [CORONAL] and [DORSAL] vowels), but also to distinguish round and unround vowels (i.e. [LABIAL] and nonlabial vowels). Thus, in a language like Dutch, where coronal/dorsal and labiality do not coincide, [LABIAL] may combine with either [CORONAL] or [DORSAL]. Rather than contrasting with [CORONAL] or [DORSAL], the presence of [LABIAL] appears to contrast with the absence of roundedness (i.e. nonlabial). This is not illustrated in a geometry such as depicted in Figure 1.5 and 1.6.

The second option was to split the node, separating [CORONAL] from [LABIAL] and [DORSAL]. For example, Rice and Avery (1995) group [DORSAL] and [LABIAL] together underneath the feature/node [PERIPHERAL] for consonants. For vowels, Rice (1995) argues that the features [DORSAL] and [LABIAL] are not even required, and there is only the feature [PERIPHERAL] for both. Rice argues that [LABIAL] and [DORSAL] do not play a role in phonology of vowels, and are treated as a single feature. This results in a system with two places of articulation: [PERIPHERAL] and placeless (i.e. not peripheral, or coronal). Based on Rice (1995), the three-way distinction between /e, ø, o/ would be represented as follows (Figure 1.7):

| /e/ | /ø/ | /o/ |
|---|---|---|
| V-Place | V-Place | V-Place |
| | / \ | \| |
| | [CORONAL] [PERIPHERAL] | [PERIPHERAL] |

**Figure 1.7** Representations of /e, ø, o/ based on Rice (1995).

Such a system grouping [DORSAL] and [LABIAL] together as a single feature seems inadequate for Dutch, because lip rounding may combine with a coronal place of articulation, whereas dorsal is mutually exclusive with coronal. It is not clear how such a feature geometry can distinguish the three mid vowels /e/, /ø/ and /o/. Since [PERIPHERAL] and [CORONAL] should be combined in case of /ø/, they should not be mapped/contrasted with each other: they are not mutually exclusive. As they share the parental node V-Place, the geometry does not reflect this. Thus, this raises an issue of mutual exclusivity, similar to no grouping of features as is assumed in the geometries depicted in Figure 1.5 and 1.6 where labial is not mutually exclusive with coronal and dorsal.

The third option proposed in the literature is to split the place of articulation node into labial and lingual features, grouping [CORONAL] and [DORSAL] together as lingual (i.e. tongue) features. As such, [LABIAL] becomes separated from [CORONAL] and [DORSAL].

From an articulatory perspective, Browman and Goldstein (1989) as well as Keyser and Stevens (1994) have proposed such geometries in which coronal and dorsal are assumed to form a single constituent – LINGUAL – because they both involve the tongue as its articulator. They argue that the lips can move independently from the tongue, but that tongue blade (= corona) and tongue body (= dorsum) are connected, and due to this anatomical connection coronal and dorsal features are not independent from each other. In contrast, the lips are anatomically and articulatorily independent from the tongue. Therefore, Keyser and Stevens assume a separate LABIAL node, whereas [CORONAL] and [DORSAL] are parented by a LINGUAL node.

From the perspective of language acquisition, similar geometries with a separate LABIAL and LINGUAL node have been proposed. Under the assumption of the building hypothesis, which assumes that children build their feature geometry by incrementally adding elements to the structure (Jakobson, 1941/68), the order of acquisition of contrasts can inform us about the feature geometry. Brown (1997) proposes a geometry separating labial from lingual features in her dissertation based on English acquisition data. The same study reported in Brown (1997) has also been published as a book chapter by Brown and Matthews (1997). Rather than unite labial and dorsal place of articulation under the feature [PERIPHERAL] as Rice and Avery (1995) do, she introduced the feature [LINGUAL], which dominates the features [DORSAL] and [CORONAL]. Brown and Matthews (1997) argue that the collected child production data (English) suggests the following order of acquisition for place features:

| | |
|---|---|
| Stage 0: | no contrast |
| Stage 1 and 2: | LABIAL VS. DORSAL followed by LABIAL VS. CORONAL - or vice versa |
| Stage 3: | CORONAL VS. DORSAL |

This suggests a split between labial and either dorsal or coronal in stage 1/2, with an added contrast between dorsal and coronal following later. Under the building hypothesis, Brown and Matthews (1997) assume this to suggest a geometry with the Place node first splitting labial and lingual features (coronal & dorsal), and the lingual feature parenting a split between dorsal and coronal. This geometry is presented in Figure 1.8.

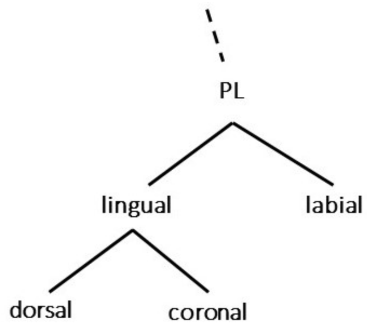



**Figure 1.8** Feature geometry as proposed in Brown (1997) and Brown and Matthews (1997). PL refers to the place node.

As far as vowels are concerned, Jakobson and Halle (1956) assumed an initial split between high /i/ and low /a/, followed by a coronal/dorsal contrast, adding /u/. Coronal labial vowels are to be added later. For Dutch vowels, Beers (1995) reports that, in production, children indeed start with the vowels /i, u, a/ - the cardinal vowels, oppositions of height, and of coronal/dorsal, and no contrast in labiality. Tense mid vowels /e/ and /o/ occur in production around the age of two. Front round vowel /ʏ/ follows around the age of three. The introduction of this vowel is the first indication in the productive inventory of children acquiring Dutch that lip rounding and coronal/dorsal are mutually autonomous: labial features no longer coincide with place features coronal and dorsal. Beers does not include tense front round vowels /y/ or /ø/ in her overview, but based on Jakobson's order of oppositions and his laws of irreversible solidarity, mid vowel /ø/ is likely to appear only after its high counterpart /y/ has become part of the system.
Levelt (1994), Fikkert and Levelt (2008), and Fikkert (2010) have argued that in child phonology the first contrast is between labial and lingual (i.e. nonlabial; default coronal). In production, Dutch children initially produce either all-labial or all-coronal words. CV-combinations share their place of articulation (PoA), similar to what has been attested in other languages (e.g. Davis, MacNeilage, and Matyear, 2002; Kern & Davis, 2009; Kern, Davis & Zink, 2010). When the vowel is a dorsal labial vowel like /o/, children tend to produce the word with labial consonants. Since this similarity is based on labiality rather than dorsality, this indicates that [LABIAL] is stored in the representation. It also indicates that at the earliest stage of acquisition, vowels and consonants are not yet separately specified and use the same features. In perception of CV-

combinations with a labial vowel, word recognition is hindered when consonants are coronal. This suggests a mismatch between perceived and stored features, thus indicating a place of articulation specification in the underlying representation. If the [LABIAL] feature of the vowel is stored to represent the entire word, [CORONAL] surface features of consonants can mismatch with the labial underlying representation. This supports early mutual exclusivity of [LABIAL] and [CORONAL], so what was considered coronal vs. dorsal supposedly is actually a contrast between [LABIAL] and [LINGUAL].

Dorsal initial words are rare in Dutch children's first words. Dorsal consonants enter production later than coronal consonants. Van der Feest, Fikkert and Davis (2016) showed that, perceptually, Dutch 24- to 26-month-olds seem to ignore the difference between coronal and dorsal stops in word-initial stops placed before coronal vowels. They report that word recognition was not hindered by place of articulation mispronunciations when the vowel following the word initial consonant was front (nonlabial), i.e. coronal. This indicates that a coronal-dorsal contrast was not yet fully established.

In summary, language acquisition data suggest an early labial vs. nonlabial contrast, with coronal as default, whereas dorsal is acquired later. When it occurs, sometimes coronals and dorsals still merge, particularly in front of a coronal vowel.

Assuming a structure where labial features are separated from lingual features predicts that coronal segments will pattern with dorsal segments in phonological operations that manipulate lingual features, but which ignore labial segments which have no lingual feature. Indeed, there are cases that support this. For example, in Mandarin Chinese, lingual obstruents are substituted by laminal palatal-alveolars before the high front vowels /i/ and /y/ (Clements, 1976). Clements and Hume (1995) also state that in Slovak, /æ/ is backed to [a] after a lingual, but not after a labial consonant. Such examples support the notion of two separate nodes for labial and lingual features, since it would explain why processes can be selective in impacting lingual, but not labial features.

In sum, I hypothesize the following feature geometry for Dutch vowels as depicted in Figure 1.9. An advantage of this geometry over the other two options discussed is the fact that this geometry reflects mutual exclusivity: the features [CORONAL] and [DORSAL] are mutually exclusive and they share their parental node LINGUAL, but they are not mutually exclusive with [LABIAL]. Since [LABIAL] may combine with either coronal or dorsal and hence does not coincide with coronal/dorsal, it should not be grouped together with [DORSAL].

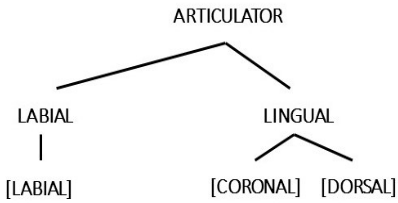



**Figure 1.9** Hypothesized feature geometry for Dutch vowels.

I assume a node parenting the feature [LABIAL], even though it parents only a single feature, because otherwise the tree would imply that [LABIAL] is contrastive with [LINGUAL], which in fact it is not. Rather, [LABIAL] is contrastive with nonlabial in vowels. However, nonlabial does not form a natural class, and therefore I do not consider this an actual phonological feature. Hence, there is a contrast between presence and absence of features below the labial node rather than a contrast between two features – like is the case below the lingual node. It follows that I assume the following features for the tense vowels of interest regarding labial and lingual features (see Table 1.3):

**Table 1.3** Assumed features of Dutch tense mid vowels.

| | /e/ | /ø/ | /o/ |
|---|---|---|---|
| [CORONAL] | √ | √ | |
| [DORSAL] | | | √ |
| [LABIAL] | | √ | √ |

In the experiments included in this dissertation in Chapter 2, 3, and 4 I will investigate whether this hypothesis holds true in perception.

Regarding /ʏ/ as the lax counterpart of /ø/, I assume lax mid vowels /ɛ/, /ʏ/, and /ɔ/, to constitute similar place and labiality contrasts as the tense vowels. These lax vowels are revisited in Chapter 3.

## Representations and processing

In this dissertation, I investigate the underlying phonology of the Dutch vowel contrasts of place and labiality. The research questions (RQ's) are:

RQ1. How is the vowel contrast between CORONAL and DORSAL (i.e. place) represented?
RQ2. How is the vowel contrast between LABIAL and nonlabial (i.e. labiality) represented?
RQ3. How do place and labiality interact when both are contrastive in vowels?

Similar to the study of German vowel contrasts by Eulitz and Lahiri (2004), I approach this from a processing perspective, collecting evidence through perception experiments conducted with Dutch listeners. However, a hypothesized feature geometry alone is not sufficient to make processing predictions. For this, one also needs to make assumptions regarding the nature of phonological representations, as well as regarding processing.

### Phonological theories

Phonological theories describe the phonological properties and patterns that surface in a language and hence need to be produced and perceived. Often these theories focus on representations. Many different phonological theories have been proposed. A lack of consensus between theories is not limited to the types of features (e.g. bivalent vs. monovalent) and their ordering in hierarchies, but also extends to the amount of detail stored in phonological lexical representations. Views regarding content of representations vary from models assuming rich phonetic detail to be part of stored representations (e.g. Johnson, 1997; Goldinger, 1998; Bybee, 2001; Pierrehumbert, 2002; Polka & Bohn, 2003, 2011; Masapollo, Polka, & Molnar, 2017; Masapollo, Polka, & Ménard, 2017) to abstractionist models that either assume full specification of features (e.g. Chomsky & Halle, 1968), or underspecified representations (e.g. Clements, 1985; Archangeli, 1988; Dresher, 2015; Lahiri & Reetz, 2002; 2010)[1].

I will focus on perception. Since word recognition is based on acoustic input, a complete theory should provide an account of how acoustics and representations are connected. There are both exemplar-based and abstractionist theories that provide such an account.

1 Pierrehumbert (2016) proposed a hybrid model which allows both episodic and abstract representations. Although Pierrehumbert is probably correct that both ends of the spectrum contain parts of the truth, a lack of detail on how both types of representations are used or interact make a hybrid model hard to test.

**Exemplar-based models** assume that the mental lexicon contains many exemplars of every word. These are all different tokens encountered by a language user, including their acoustic/phonetic detail. Together, these exemplars form a word cloud (Ernestus, 2014). Hence, words are stored with a lot of phonetic and redundant detail. Exemplar-based models treat surface variance in pronunciation by having a representation for all surface variants. Stored representations do not abstract away from contextual speaker or situation specific characteristics (Ernestus, 2014). Rather, a basic assumption is that underlying representations are episodic in nature.

Within exemplar-based models word recognition is assumed to occur through direct mapping of the perceived acoustic input onto the exemplars in the mental lexicon. An exemplar is activated depending on how well it matches the acoustic input, and it passes its activation to its word node. The word node receiving the most activation is then selected as the recognized word. Word recognition occurs without a mediating pre-lexical representation, and without speaker normalization. Such a model hence does not explain word recognition in terms of phonological features.

Rather than a cloud of exemplars with phonetic detail, **abstractionists** assume that the mental lexicon contains only a single underlying representation for every word. Abstractionists assume that during word recognition one disregards speaker and context/situation specific characteristics. The underlying representation consists of abstract symbols represented by phonological features[2]. This view started with Jakobson, Fant and Halle (1952) who presented a provisional list of distinctive features, followed by The Sound Pattern of English (SPE) by Chomsky and Halle (1968) - which has inspired many others.

Abstractionist models vary considerably in which features they assume are stored in the underlying representation. **Fully specified models** assume that all phonological features are stored in the underlying representation, whereas **underspecified models** assume not all features need to be stored. Some features may be redundant or predictable based on other features, and therefore need not be specified (e.g. Archangeli, 1988; Keating, 1988; Steriade, 1995). An extreme version of underspecification is the Featurally Underspecified Lexicon (FUL) model (Lahiri & Reetz, 2002, 2010; Lahiri, 2018), already introduced in the previous section (Feature geometry).This model assumes that only contrastive information is stored, whereas not only predictable information, but also features that could be called unmarked or default may be underspecified. Unlike for example Dresher's account of underspecification (e.g. 2005; 2015), which assumes language specific markedness and geometry, FUL aims to define and regulate a set of features which can cover the typology of all possible contrasts in the languages of the world (Lahiri, 2018). Most underspecification accounts are based on

2 This does not mean that other characteristics are not stored, they simply are not part of the lexical representation.

production to explain asymmetries in assimilation processes, such as nasal assimilation of place of articulation: a coronal /n/ may be produced as a labial [m] when followed by a labial obstruent (e.g. *garde[m]bench)*, whereas an /m/ does not become an [n] when it is followed by a coronal obstruent.

In sum, numerous choices can and have been made in different models when it comes to phonological theories on features and their representations. However, word recognition cannot be explained based on representations only. An account of how representations are used to get from acoustic input to word recognition is required. The FUL model is a present-day phonological model that combines both representations and processing (e.g. Lahiri & Reetz, 2002; 2010; Kotzor et al., 2017; Lahiri, 2018).

The FUL model makes the following assumptions regarding features and geometry: (1) All features are monovalent/privative (i.e. presence or absence of features). (2) They are placed in a universal feature tree. Regardless of assuming a specific geometry, (3) FUL assumes no dependencies. No feature is ever subordinate to another feature, but rather they are subordinate to nodes. Features are always end points. (4) FUL assumes features under the same node to be mutually exclusive. (5) The same features and geometry are assumed for both consonants and vowels (i.e. unified features).

FUL makes very specific predictions for perception based on representations. There are two levels of representations: surface representation and underlying representation. The model assumes that listeners extract features from the acoustic signal. These extracted features are called surface representations (SR). Surface features are mapped onto the features of the underlying representations (UR) in the mental lexicon. A feature is either *specified*, *underspecified* or *not specified* at all:

1. ***Specified*;** the feature is stored in UR. This is the case for marked distinctive features which have acoustic correlates. They hence also exist at SR level.
    - For example, when a sound is [DORSAL], this feature is stored in UR, and it is also extracted from the acoustics when a dorsal sound is heard and translated into an SR feature.
2. ***Underspecified*;** the feature is not stored in UR, but it is filled in during articulation. This occurs when a feature is considered default/unmarked or redundant/predictable. Underspecified features do have acoustic correlates and may hence be extracted and exist at SR level.
    - For example, the feature [CORONAL] is underspecified for a coronal sound. This leads to an asymmetry between surface representation and

underlying representation: [CORONAL] may be present at SR, but not in UR, for which we use the symbol [Φ].

3. ***Not specified*;** information for the feature is absent in UR - the representation is *empty* altogether. No feature is extracted from the input either. This is the case when a feature is non-distinctive and/or not perceivable: it lacks an acoustic cue and/or is not a natural class.
    - For example, voicelessness may be considered *not specified*. It does not form a natural class (determined by participation in shared phonological processes), nor does it have a clear acoustic cue (Lahiri, 2018)

In summary, a feature should be distinctive to be either specified or underspecified. Also, it needs to have a clear acoustic cue, in order to be extractable from the signal (SR). Both specified and underspecified features exist at SR level. Specified features are considered marked. Underspecified features are considered default or redundant. For underspecified features the underlying representation lacks specific information about the feature, but this feature is filled in with a default value in articulation. FUL assumes the features [CORONAL] and [PLOSIVE] to be underspecified in all languages. Importantly, a feature's parental node(s) only exist(s) when one of its daughter features is present (either specified or underspecified).

A distinction is made between features being *underspecified* and features being *not specified* (Lahiri & Reetz, 2010; Lahiri, 2018). Unlike what is the case for underspecified features, features that are *not specified* are not filled in production (Lahiri & Reetz, 2010). A feature may be *not specified* (i.e. *empty*) in case it is not distinctive in the particular phonological system or in case it is not possible to find an acoustic correlate. An example given by Lahiri and Reetz (2010) regards the voicing contrast. In that case, the feature [VOICE] is specified, but because voicelessness does not have an acoustic cue, it does not exist as a feature in FUL. Therefore, consonants which are not voiced are *not specified* in production, and voicelessness cannot be extracted from the input by the perceptual system. Thus, voicelessness is neither part of SR nor of UR. The impact of this on perception of a contrast between a specified and a not specified (and hence non-existent) feature will be revisited in Chapter 5.

A ternary mapping procedure is employed to find the parsimonious candidate in word recognition; the best match receives the highest level of activation[3]. Features from the input (SR) are compared to stored features (UR). If a phonological feature is extracted, mapping can result in three different outcomes:

3 There may be multiple lexical entries activated based on the mapping system – FUL over-generates options. Ultimately, for word recognition, other factors like expectations/context are taken into account in deciding on which word was heard.

1. ***Match**;* occurs when an SR feature matches the feature stored in the UR; for example, a dorsal feature from the surface will match a dorsal feature in UR. This will strengthen activation of a word candidate.
2. ***Mismatch**;* occurs when an SR feature is conflicting with the one stored in the UR; i.e. mutually exclusive. For instance, when a [CORONAL] feature is mapped onto a dorsal feature in UR, this will result in a mismatch. This will eliminate the word candidate.
3. ***No-mismatch**;* occurs when the SR feature neither matches nor mismatches with the UR. This occurs when features are underspecified. For example, FUL assumes the feature [CORONAL] to be underspecified. Mapping a dorsal SR feature onto an underspecified coronal feature will result in a no-mismatch, i.e. no phonological conflict. As such, a word candidate is still considered a potential target word.

FUL's ternary mapping procedure predicts specific asymmetries in perception; as no surface feature can mismatch with underspecified features, underspecified features allow for more surface variation than specified features. As a result, some featural changes are more readily detected in perception than others. More concretely, this means that a change from one phoneme into another phoneme may be more prominent than vice versa (see see the next section on perceptual asymmetries for examples). As such, underspecification not only explains asymmetries in production processes, but also predicts asymmetries in perception. Full specification on the other hand, predicts symmetry, since all features are stored and hence all featural differences between surface form and stored form will lead to conflict in perception.

The above illustrates how assumptions regarding the nature of representations and mapping of surface and underlying representations have processing consequences. For example, full specification, would not predict asymmetries in perception whereas underspecification would. What I mean by asymmetries in perception will become more concrete in the next section.

**Perceptual asymmetries**

Experimental evidence of asymmetries in perception shows that Dutch toddlers detect a mispronunciation in a preferential looking paradigm when labial /p/ is mispronounced as coronal [t], but not vice versa (Van der Feest & Fikkert, 2015). For example, when */p/oes* ('cat') is mispronounced as **[t]oes*, children's recognition of the picture of a cat is hindered compared to the correct pronunciation, which means that they successfully detect the difference. However, vice versa, when the coronal stop in */t/and* ('tooth') is mispronounced as labial *[p]and*, this mispronunciation does not hinder word recognition. The time spent looking at a picture of a tooth did not differ between hearing the correct and mispronounced word (Van der Feest, 2007; Van der Feest & Fikkert, 2015). Within a single contrast (in this case a coronal/labial contrast in consonants), opposite mispronunciations comprising identical acoustic distances result in different perceptual distances depending on the direction of the mispronunciation. In line with FUL, Van der Feest and Fikkert assume /t/ to be underspecified for [CORONAL]. Therefore, mapping of the [LABIAL] surface feature of its mispronunciation [p] results in a no-mismatch. Therefore, [p] is accepted as a variant of /t/. However, because [LABIAL] is stored in UR for /p/, a [CORONAL] mispronunciation [t] results in a mismatch. Thus, [t] is not accepted as a surface variant of /p/.

Experiments in adult vowel perception also show asymmetries that can be predicted/ explained using the same logic. Neurobiological evidence from electroencephalographic (EEG) experiments shows that for both German and French adult listeners a change from [DORSAL] /o/ to [CORONAL] [ø] yields larger electrical brain responses, indicating a larger perceptual change, than vice versa (Eulitz & Lahiri, 2004; De Jonge & Boersma, 2015), despite both changes constituting an identical absolute acoustic distance. Both experiments measured the electrical brain response to a vowel change and compared responses for opposite changes within a contrast. Again, authors interpret this as coronal underspecification; As coronal /ø/ is underspecified for place, a [DORSAL] surface feature of [o] will not mismatch the underlying representation of /ø/, because there is nothing to mismatch with. Thus, no conflict is detected and /ø/ is still considered a possible target, leading to a rather small perceptual difference and a small brain response. Vice versa, the [CORONAL] SR feature of [ø] will mismatch with the stored [DORSAL] representation of /o/. The large perceptual distance due to this mismatch is reflected by a large brain response for this condition.

Many more perceptual asymmetries have been found in a variety of languages and phonological contrasts – in words as well as in phonemes in isolation (e.g. Eulitz & Lahiri, 2004 (coronal-dorsal contrast in German); De Jonge & Boersma, 2015 (coronal-dorsal and height in French); Roberts, Kotzor, Wetterlin & Lahiri, 2014 (consonant duration in Bengali); Kotzor, Roberts, Wetterlin & Lahiri, 2015 (nasality in Bengali); Scharinger, Monahan & Idsardi, 2012 (height in English); Scharinger & Lahiri, 2010 (height in different English dialects), Scharinger et al. (2010) (coronal-dorsal in German words); Van der Feest & Fikkert, 2015 (voicing and coronal-labial in consonants in Dutch toddlers); Altvater-Mackensen, Van der Feest & Fikkert, 2014 (stop-fricative); Dijkstra & Fikkert, 2011, and Tsuji, Fikkert, Yamane & Mazuka, 2016 (coronal-labial word-initial consonants)).

Since phonological asymmetries are prevalent, phonological theory should be able to account for them. Asymmetries in perception suggest that in one direction there is a mismatch/conflict and in the other there is not. Models that assume full specification do not predict asymmetries: when all features are stored, all discrepancies between the representation and the surface form will result in conflict regardless of direction of change. While full specification is still the norm in psycholinguistic research, underspecified representations may better account for perceptual asymmetries.

The FUL model accounts for asymmetries at a phonological level due to its notion of underspecification, in combination with its ternary mapping procedure. The FUL model provides powerful predictions for perception, is widely tested, and has proved promising. The FUL model aims to define and regulate a set of features which can cover the typology of all possible contrasts and alternations in the languages of the world (Lahiri, 2018). In addition, it is able to account for acquisition and language processing. In perception, processing demands on computing power are relatively low, allowing language processing to be efficient and swift. The FUL model has a strong ecological validity, overarching both perception and production, and focusing on the language user.

Since FUL necessarily predicts asymmetries in perception, investigating the presence and direction of such asymmetries can provide insight into the underlying phonology. Perceived phonological conflict, or lack thereof, can inform us on what is stored in the phonological representations, and how these representations are used during word recognition. I regard the FUL model an advantageous starting point for investigating phonological representations of the Dutch vowels.

In the FUL model, claims on representations (phonology) can be viewed independently of the assumed mapping procedure (processing). The FUL model thus elegantly combines the field of phonology and the field of psycholinguistics. This combination allows to test assumptions on phonological representations in perception. Also, FUL's featural representations and mapping procedure can be viewed independently from the feature geometry. In the current thesis, I hypothesize a different geometry than the FUL model for the reasons presented in the section Feature Geometry, but I adopt FUL's logic regarding underspecification and mapping of features in order to make predictions for perception of the labiality and place contrasts in Dutch vowels.

## Predictions

Dutch listeners need to keep the coronal-dorsal contrast and the contrast between labial and nonlabial (=unround) apart, since LABIAL is not mutually exclusive with CORONAL or DORSAL, but rather is contrastive with nonlabial. It is not evident how the Dutch three-way phonological contrast is perceived. In this thesis, I collect perceptual data of the Dutch vowel contrasts of labiality and place. In particular, I investigate perception of the contrasts between /e/, /ø/ and /o/. Below I provide predictions based on exemplar-based models, full specification, FUL's underspecification account, and my own account combining FUL's logic with the feature geometry I argued for.

Experience-based theories may explain asymmetries in perception in terms of frequency or saliency. For frequency, the **Native Language Magnet (NLM) model** (e.g. Kuhl, 1991; Kuhl et al., 2008) assumes that category building is based on distribution of input, and more frequent vowels have a stronger magnetic effect, warping the perceptual space around them. Based on the most frequently heard forms within a category (or cloud of exemplars) a prototype is established, which functions as a magnet in vowel perception. Poorer discrimination is expected in the direction from more frequent to less frequent. However, this discrimination effect usually holds for within-category discrimination. For between-category discrimination, the topic in this dissertation, predictions are less clear. One could hypothesize that frequent vowels accept more variation since the prototype is a stronger magnet to other (deviating) vowels. Then, it is expected that a change from a more frequent to a less frequent vowel would be more difficult to detect than vice versa. Baayen, Piepenbrock, & Gulikers (1995) report the following frequencies for the relevant vowels in CELEX: [e] = 6,7%; [o] = 6,0%; [ø] = 2,3%, meaning /e/ and /o/ occur twice as often as /ø/. Dutch front round vowels are relatively infrequent compared to front (not round) vowels and back (round) vowels. A change from either /e/ or /o/ toward /ø/ is thus expected to be more prominent than vice versa. Predictions are listed in Table 1.4.

On the other hand, the **Natural Referent Vowel (NRV) framework** (Polka & Bohn, 2003; 2011; Masapollo, Polka, & Molnar, 2017; Masapollo, Polka, & Ménard, 2017) would explain asymmetries based on acoustic salience. It assumes that vowels with formant frequencies closer together have focalized energy, and hence, are universally more salient in perception than vowels with formants further apart. The most focalized vowels are /i/, /a/, and /u/, the cornerstones of vowel space. Consequently, changes from less to more focal vowels are easier to discriminate. Based on the convergence or closeness of second and third formant (F2-F3), /e/ is more focal than /ø/. Based on the first and second formant (F1-F2), /ø/ would be more focal than /e/. In other words, the predictions based on the NRV are not entirely straightforward. For the vowel pair /o/-/ø/ the predictions are not clear either. Based on F1-F2, /o/ is more focal than /ø/. Based on F2-F3 /ø/ is more focal than /o/ - see Table 1.4.

As mentioned in the previous section, **full specification** of abstract features does not predict any asymmetries, because all features are stored and hence all featural differences between surface form and stored form will lead to conflict in perception. Thus, also for the vowel contrasts tested in this thesis, no asymmetries are predicted based on full specification.

**The FUL model** would in fact predict a coronal-dorsal asymmetry for the contrast between [DORSAL] /o/ and [CORONAL] /ø/, where [LABIAL] is held constant. FUL assumes [CORONAL] to be universally underspecified. Dorsal is distinctive and hence it would predict that dorsal is in fact stored. This would mean that mapping of SR of a [DORSAL] [o] on UR of /ø/ would result in a no-mismatch, which means that no conflict between SR and UR is detected. Vice versa, mapping of the extracted feature [CORONAL] of [ø] onto the UR of [DORSAL] /o/ would result in a mismatch. Thus, based on FUL, a coronal-dorsal asymmetry would be predicted for the place contrast between /o/ and /ø/.

Predictions based on FUL for the labiality contrast between /e/ and /ø/ are not clear. As discussed in previously, FUL's geometry implies that the feature labial would be mutually exclusive with coronal and dorsal. This would imply that the feature labial would be mapped onto either of those other features. However, this would not be a meaningful comparison as labial may be combined with either coronal or dorsal. Rather, it is contrastive with the absence of labiality. Although Lahiri (2018) does state an exception for the mutual exclusivity assumption with respect to vowels, there is no explanation of how this works. Hence, it remains unclear how FUL would explain the discrimination of a contrast between vowels that merely differ with respect to labiality at a phonological level.

As a step in resolving this, I hypothesize a geometry where the feature labial is separated from the lingual features coronal and dorsal (see Figure 1.9). This does not change predictions for the coronal-dorsal contrast between /ø/ and /o/ compared to FUL. Coronal and dorsal still share their parental node. Assuming coronal underspecified and dorsal specified, similar mapping outcomes are expected as discussed for FUL. The different geometry explains why the feature [LABIAL] is now not mapped onto [CORONAL] or [DORSAL]. This enables us to make clear predictions for the labiality contrast. No asymmetry is predicted for the labiality contrast between *nonlabial* /e/ and [LABIAL] /ø/, where [CORONAL] is held constant. FUL assumes [LABIAL] to be specified in the underlying representation of /ø/, and it predicts no specification for labiality for /e/. When SR of [e] is mapped onto UR of /ø/, no mapping of labiality features takes place, simply because no labiality information is extracted from the signal, which is necessary to initiate mapping. Mapping of SR features of [e] onto UR features of /ø/ will thus not result in a mismatch and no phonological conflict will be perceived. Vice versa, when [ø] is heard, [LABIAL] can be extracted and mapped to the UR of /e/. As /e/ is not specified for labiality, mapping will not result in a mismatch either. Accordingly, similar degrees of conflict for both directions of change are predicted – see Table 1.4.

**Table 1.4** Predictions based on exemplar-based and abstractionist accounts regarding place and labiality vowel contrasts between /e/, /o/ and /ø/.

| | models | | place | labiality |
|---|---|---|---|---|
| **exemplar-based** | **frequency (NLM)** | | /ø/ → /o/ | /e/ ← /ø/ |
| | **saliency (NRV)** | F1-F2 | /ø/ → /o/ | /e/ → /ø/ |
| | | F2-F3 | /ø/ ← /o/ | /e/ ← /ø/ |
| **abstract** | **full specification** | | /ø/ - /o/ | /e/ - /ø/ |
| | **FUL model** | | /ø/ ? /o/ | /e/ → /ø/ |
| | **current hypothesis** | | /ø/ - /o/ | /e/ → /ø/ |

Arrows indicate the predicted direction of change which is most easily detected in perception. – indicates symmetry. ? indicates an unclear prediction.

As discussed previously, a phonological system that results in no phonological conflict in either direction of change in a contrast that is in fact important in a language seems inadequate. Thus, the hypothesized model still raises questions. Experimental data collected in this thesis can provide more insight into how the labiality contrast is accounted for in Dutch.

In addition to the single-feature contrasts of place (/ø-o/) and labiality (/e-ø/), the two-feature contrast between coronal /e/ and dorsal (labial) /o/ is investigated. These two vowels differ with respect to both labiality and place. For this contrast, predictions based on exemplar-based accounts are unclear. It is not clear what NLM would predict, as predictions would be based on differences in frequency of occurrence, but /e/ and /o/ are rather similar in frequency. Also, the NRV framework would not make clear predictions as both vowels are considered peripheral and hence focal, albeit one based on F1-F2 and the other based on F2-F3. It is not clear what this framework would predict for this case.

From an abstractionist account, full specification would predict phonological conflict in either direction of change for the contrast between /e/ and /o/. No asymmetry would be predicted.

From an underspecification account, FUL with its own geometry as well as my current hypothesis with a different geometry both predict that SR of coronal /e/ would mismatch with UR of dorsal /o/. If indeed labiality does not trigger any mismatches (as explained for the contrast between /e/ and /ø/), a coronal-dorsal asymmetry would be predicted for the contrast between /e/ and /o/, similar to what is predicted for the contrast between /ø/ and /o/. However, outcomes may be different. After all, /e/ and /o/ are acoustically further apart than the single-feature contrasts. Eulitz and Lahiri (2004) suggest that this may be why they did not find a coronal-dorsal asymmetry for the German /e-o/ contrast. Also, it has been suggested in the literature that with featural distances larger than one feature, the underspecification interpretation as provided by FUL may not hold (Scharinger, Domahs, Klein & Domahs, 2016); Scharinger and colleagues (2016) observed that a change from less to a more specified sound elicited stronger responses than vice versa, rather than the other way around. Lastly, the currently hypothesized phonological system still seems to fail to explain how the labiality contrast is discriminated. Labiality appears to play an important role, and this may be reflected in processing. Thus, it remains speculative whether the prediction of the place contrast will drive a perceptual asymmetry between /e/ and /o/.

## Outline of the thesis

This dissertation includes three chapters describing perception experiments that investigate the underlying phonology of the Dutch vowel contrasts of place and labiality in both adults and children (Chapter 2, 3 and 4). Each experiment focuses on the same three research questions (RQs), here recapitulated:

RQ1. How is the vowel contrast between CORONAL and DORSAL (i.e. place) represented?

RQ2. How is the vowel contrast between LABIAL and nonlabial (i.e. labiality) represented?

RQ3. How do place and labiality interact when both are contrastive in vowels?

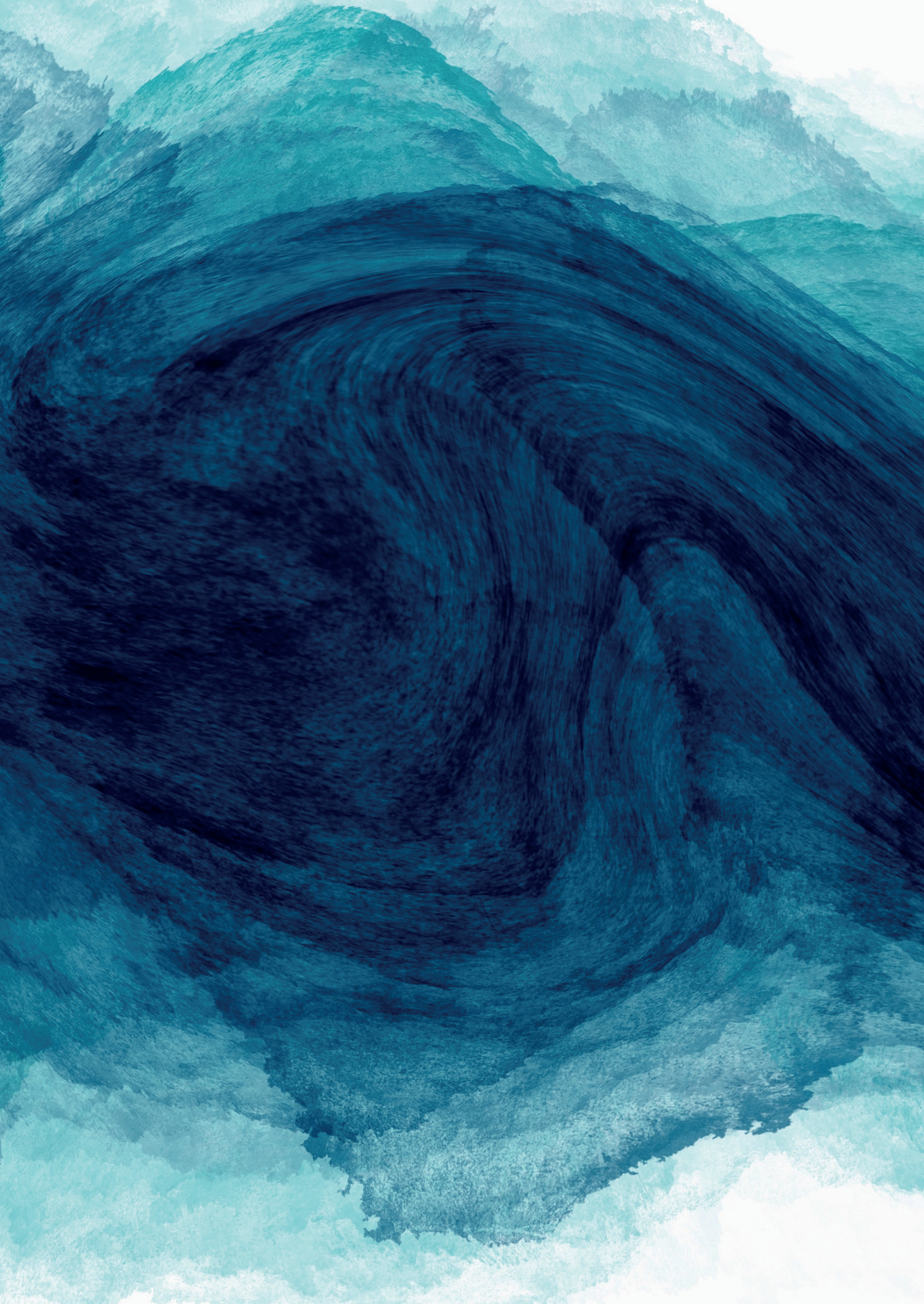
Chapter 2 and Chapter 3 discuss perception experiments in adults. In **Chapter 2**, the EEG experiment as performed in German by Eulitz and Lahiri (2004) is replicated in Dutch, testing perceptual difference between the three tense mid vowels /e/, /o/ and /ø/, presented as vowels in isolation. Electrical brain activity in response to vowel changes is measured using a passive oddball paradigm. The mismatch negativity (MMN) component – an automatic change detection response of the brain – is used as a proxy of perceptual difference for both directions of change within the contrasts of place (/o-ø/; RQ1), labiality (/e-ø/; RQ2), or both (/e-o/; RQ3). Data is tested on asymmetries, providing insight into the underlying phonology of contrasts of place and labiality. The language systems of German and Dutch differ in important ways, so similar outcomes are not evident. German findings (Eulitz & Lahiri, 2004) and Dutch results are compared. **Chapter 3** elaborates on this EEG experiment by means of a lexical decision experiment with semantic priming. In this experiment, I investigated whether or not mispronunciations of vowels regarding place and/or labiality features hinder word recognition, by measuring reaction times rather than electrical brain activity. This experiment investigates whether findings of Chapter 2 are replicated.

**Chapter 4** focuses on children's phonological representations. Investigating child phonology is valuable, because whereas adults have a complex and complete system that may be hard to disentangle, children acquiring their mother tongue are building up their system step by step.

As such, children may not necessarily have all contrasts yet, or they may not be fully engrained. The child system provides insight into how the adult system is built, and as a consequence can inform us on what is possible for the adult system. As such, child phonology can impact phonological theory.

In the experiment discussed in Chapter 4, word recognition of three-year-olds is tested using a preferential looking paradigm with place and labiality mispronunciations of vowels /e, o, ø/. Front round vowels are typically acquired relatively late (e.g. Beers, 1995). Three-year-old Dutch toddlers have a limited number of words with front rounded vowels in their lexicon (N-CDI / N-CDI-3; Zink & Lejaegere 2002; 2007). They only just started producing words with front rounded vowels (e.g. Beers, 1995), which makes this an interesting age to investigate whether or not they are already sensitive to the three-way distinction between /e, o, ø/ at a lexical level, and how their phonological system accounts for contrasts of place and labiality in vowels. Do they show the same patterns in perception as the adults do?

The main findings of these three experiments are summarized and discussed in **Chapter 5**. This chapter also presents the key conclusions and suggests directions for further research.



# Chapter 2

## Asymmetric processing of place and labiality: Electrophysiological evidence

Based on: de Rue, N.P.W.D., Snijders, T.M., & Fikkert, P. (2021).
Contrast and conflict in Dutch vowels.
*Frontiers in Human Neuroscience*, 15:629648.
https://doi.org/10.3389/fnhum.2021.629648

## Introduction

There is considerable acoustic variation in natural speech, making recognition of spoken words or even single vowels rather complex. For word recognition, it is important to perceive meaningful differences (i.e. phonological features), which are contrastive and to some extent predictable. The phonological underlying representations of words and phonemes in our mental lexicon, are made up of contrastive features and they in turn play a role in recognition. A vital issue in phonology is what information representations contain, and how they enable us to recognize and produce words the way we do. Some theories assume rich phonetic detail to be part of stored representations (e.g. Johnson, 1997; Goldinger, 1998; Bybee, 2001; Pierrehumbert, 2002; Polka & Bohn, 2003, Masapollo, Polka, & Molnar, 2017; Masapollo, Polka, & Ménard, 2017), whilst others only assume the essential features needed to differentiate between lexical contrasts (e.g. Chomsky & Halle, 1968, Archangeli, 1988; Lahiri & Reetz, 2002; Dresher, 2009). For example, the Featurally Underspecified Lexicon (FUL) model assumes underspecification of monovalent phonological features (i.e. presence or absence), and assumes only contrastive features to be stored in mental representations.

Dutch vowels are of interest as they have a contrast between labial and nonlabial vowels, as illustrated by the three-way contrast between coronal (i.e. front) /e/, coronal labial (i.e. front round) /ø/ and dorsal labial (i.e. back round) /o/. This is unlike most languages, which contrast front nonlabial and back labial vowels, hence either place or labiality would suffice to discriminate these vowels. In Dutch, labiality does not coincide with place.

As introduced in Chapter 1, the current dissertation focuses on how the place and labiality contrasts are perceived in Dutch. We use the FUL model's account of representations and processing to make predictions (see Chapter 1 for more detailed information about the FUL model). In FUL, redundant or default features may be underspecified. In particular, FUL assumes the feature coronal to be universally underspecified. Furthermore, the model assumes a ternary mapping procedure. Listeners extract features from the acoustic signal which are mapped onto the features in the underlying representations in the mental lexicon. The perceived surface features can *match*, *mismatch* or *no-mismatch* with the underlying features. Matches and mismatches require features to be detected from the input and present in the underlying representation. A no-mismatch occurs when an extracted surface feature is underspecified in the underlying representation (e.g. default), so there is nothing to (mis)match with. A mismatch indicates a phonological conflict, whereas a match or a no-mismatch do not.

Although we follow these assumptions in FUL, we assume a different feature geometry (argued for in Chapter 1). Rather than [CORONAL], [DORSAL] and [LABIAL] sharing their parental node, we hypothesize [CORONAL] and [DORSAL] to be separated from [LABIAL], as illustrated in Figure 2.1.

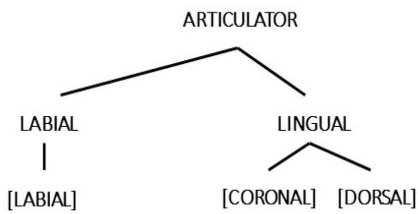



**Figure 2.1** Hypothesized feature geometry.

At the end of the current section, we will formulate predictions for perception of the contrasts of interest. In the current chapter, these predictions are tested during a perception experiment, measuring electrical brain activity in response to vowel changes.

**Asymmetries in EEG**

As explained in Chapter 1, in perception, a change from one phoneme to another may be more prominent or easy to detect than an opposite change, even though the acoustic distance is identical. Perceptual asymmetries in vowels have been found in a variety of languages and phonological contrasts – in words as well as in speech sounds in isolation using a variety of methods (e.g. Eulitz & Lahiri, 2004 (vowel place and labiality in German); De Jonge & Boersma, 2015 (vowel place and height in French); Roberts et al., 2014 (duration in Bengali); Kotzor et al., 2015 (nasality in Bengali); Scharinger et al., 2012 (height in English); Scharinger & Lahiri, 2010 (height in different English dialects), Scharinger et al. (2010) (vowel place in German words)) as well as in words and nonwords (Cornell et al. (2011): vowel place in German). In these studies, authors argued that the reported asymmetries help us understand which phonological features can be assumed stored or not.

Perceptual asymmetries can be attested by directly measuring electrical brain activity using electroencephalography (EEG). The Mismatch Negativity (MMN) response has been shown to be sensitive to language specific phoneme representations or phoneme boundaries (e.g. Phillips et al., 2000; Winkler et al., 1999; Näätänen et al., 1997; Näätänen, 2001 & 2007; Dehaene-Lambertz, 1997). It is elicited when in a stream of standard (or frequent) sounds suddenly a deviant (or infrequent) sound is presented (oddball paradigm). The standard stimuli create a central sound representation that is more abstract than the sum of perceived acoustic elements (Näätänen, 2001). Apart from sound perception and sensory memory, this representation corresponds to long-term memory (Näätänen, 2001; Cowan, 1999). That means that the central sound representation may convey information about the phonological representation in the mental lexicon, which is the underlying representation. The repetitive nature of the standard stream of sounds creates an expectation that the next sound will be part of the same phonemic category as well, and because of this expectation, the surface representation of the deviant sound is mapped onto the underlying representation of the standard sound. The percept created by the deviant stimulus is more low level (Eulitz & Lahiri, 2004). Vowel specific information is available around 100 ms after stimulus onset (e.g. Eulitz & Lahiri, 2004; Obleser & Eulitz, 2002; Obleser, Elbert, Lahiri, & Eulitz, 2003; Poeppel et al., 1996; Eulitz, Diesch, Pantev, Hampson & Elbert, 1995). Apart from sound perception (i.e. acoustics), this percept corresponds to the phonological features extracted from the acoustic signal, the so-called surface form or SR (surface representation). A deviant sound violates the expectation and elicits a mismatch response. If mapping of SR and UR features results in a phonological mismatch, the perceptual difference between standard and deviant is enhanced compared to when they are not conflicting. This is reflected by the MMN: larger perceptual differences are reflected by MMNs with a larger amplitude and/or earlier latency, whereas smaller perceptual differences result in smaller amplitudes and/or later peaks. Although MMN strictly reflects neural discrimination, this is commonly used as a proxy for perception.

The MMN response can be understood to reflect the outcome of the mapping procedure: As illustrated in, for example, the paper by Eulitz and Lahiri (2004), an enhanced (large and/or early) MMN response is interpreted as a mismatch, and hence implies phonological conflict. When mapping results in a no-mismatch, which implies no phonological conflict, this is reflected by a smaller/later MMN. In summary, asymmetries in MMN properties can inform us about the degree of conflict and subsequently also on the representations.

**Place and labiality in German and Dutch vowels**

In their seminal MMN study, Eulitz and Lahiri (2004) showed that German listeners perceived native vowel contrasts asymmetrically, and argued that this asymmetry is due to phonological underspecification of coronal place of articulation. While the acoustic difference between mid vowels /e, ø/ on the one hand, and /ø, o/ on the other is similar, their phonological representations differ (see Figure 2.2).

The contrast /e, ø/ shares the same place of articulation [CORONAL], but differs in the value for [LABIAL]: /ø/ is [LABIAL], while /e/ is not. The contrast /ø, o/ differs in place of articulation, /ø/ being [CORONAL], while /o/ is [DORSAL], but they share the feature [LABIAL]. While /e, ø/ are underspecified for place of articulation in the underlying representation, indicated by [––] in Figure 2.2, /o/ is specified as [DORSAL]. Furthermore, /ø, o/ both are specified for [LABIAL]. However, the underspecified feature [CORONAL] can still be extracted from the acoustic signal, and thus be part of the surface representation, just like [LABIAL] and [DORSAL]. When the standard is presented in an oddball paradigm, the underlying representation of the standard is preactivated. Upon hearing a deviant, the surface representation of the deviant is mapped onto the preactivated underlying representation of the standard. When the standard is /o/, the coronal place of articulation of deviant [ø] will lead to a mismatch (marked with a red line), but the reverse will not, as the perceived [DORSAL] feature of [o] will form a no-mismatch with the underspecified [CORONAL] in the underlying representation of /ø/, as shown in Figure 2.2.

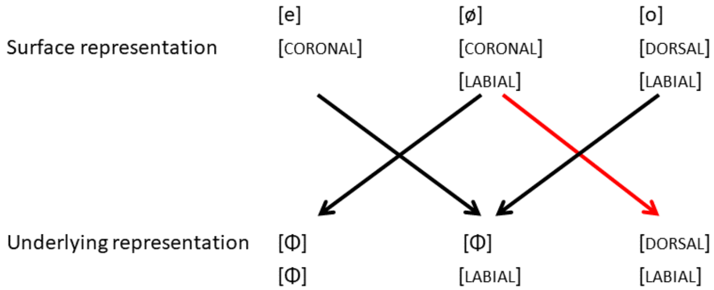



**Figure 2.2** Representations and mapping outcomes as given for German by Eulitz and Lahiri (2004) based on FUL. Red arrow indicates a mismatch, i.e. phonological conflict. Black arrows indicate no conflict. [Φ] indicates underspecification.

Based on this perceptual asymmetry, Eulitz and Lahiri (2004) argue that [CORONAL] is underspecified in German. The perceptual asymmetry of coronal and dorsal have also been found in subsequent studies for vowels, both in words and nonwords (Cornell et al., 2011). The role of [LABIAL] is not much discussed in the paper by Eulitz and Lahiri (2004).

Like German, Dutch has a three-way contrast between /e, ø, o/. Yet, looking at the linguistic system as a whole, the nature of the front rounded vowels may be different in Dutch and German.

First, front round vowels do frequently occur in German, but are typically derived from back round vowels due to morphological umlaut (i.e. fronting) in, for example, plurals or diminutives. For example, German has the singular-plural alternation *V/o/gel* – *V[ø]gel* 'bird(s)', whereas Dutch has *vogel* – *vogels*, without vowel alternation. The original motivation for umlaut was phonological (Twaddell, 1938), but in modern German umlaut is no longer phonologically transparent. Rather, it has become a morphological rule (Wiese, 1996). Consequently, the /ø/ predominantly occurs in morphologically derived contexts, and might not be part of the underlying vowel inventory of German. As a result, the place contrast between front round and back round vowels signals morphological rather than lexical significance in German. Scharinger (2009) has argued that the underlying representations of stems that may alternate between front round and back round vowels (due to umlaut) in morphologically derived context may differ from stems that do not alternate. Non-alternating stems with front round vowels are far less frequent in German than in Dutch. Modern Dutch does not have morphological umlaut, and front rounded vowels are truly part of the Dutch vowel inventory. Front round vowels are not derived, but rather signal lexical contrast with both front unround and back round vowels.

Second, in German, the contrast between labial and nonlabial vowels is of minor importance compared to Dutch. Minimal pairs of word stems that merely differ regarding labiality of a vowel are extremely rare in German. If two words differ regarding labiality of the vowel, they typically also differ regarding one or more other phonological features. Thus, distinguishing between labial and nonlabial vowels may not be required in German in order to learn and recognize all words. In contrast, many word stems in Dutch differ merely with respect to labiality of a vowel, e.g. st/ø/n (= support) vs. st/e/n (= rock). Front round vowels are much more eminent in Dutch monomorphemic words than in German. The labiality contrast plays a substantial lexical role in Dutch monosyllabic words. Dutch listeners therefore need to resolve the contrast at a phonological level in order to learn and recognize all words.

Despite front round vowels being frequent in both German and Dutch, labiality and place contrasts in vowels play a different role in both languages. We investigate whether this difference between German and Dutch is reflected in differences in phonological processing.

**Current study**

In the current study we investigate the processing of a three-way vowel contrast in Dutch. Similar to Eulitz and Lahiri (2004), we measured the MMN in an EEG experiment, but with Dutch listeners. We tested discrimination of tense mid vowels /e/, /o/ and /ø/ in isolation in a passive oddball paradigm. The three vowels /e/, /o/ and /ø/ form single-feature contrasts with respect to place (/ø,o/), as well as labiality (/e,ø/), but also provide a two-feature contrast where both place and labiality are contrastive (/e,o/).

Predictions are based on our hypothesized geometry - which separates [LABIAL] from [CORONAL] and [DORSAL] - in combination with FUL's logic regarding mapping procedure and underspecification. Like Eulitz and Lahiri (2004), we hypothesize that the MMN would be enhanced (amplitude and/or latency) if mapping of deviant SR on standard UR constitutes a conflict rather than a non-conflict situation (see also Näätänen & Alho, 1997). Below we describe our predicted outcomes for each tested contrast. An overview is given in Table 2.1 (next page).

We test the place contrast between [DORSAL] /o/ and [CORONAL] /ø/, where [LABIAL] is held constant. We predict to find a perceptual asymmetry for place, supporting the underspecification of [CORONAL] and replicating the results for German. We expect [CORONAL] to be underspecified and [DORSAL] to be specified in UR. This would mean that SR of a dorsal [o] as a deviant would result in a no-mismatch with the UR of standard /ø/, which means that no conflict between SR and UR is detected. This is expected to result in a rather small/late MMN response. Vice versa, SR of coronal [ø] as a deviant would mismatch with underlying [DORSAL] representation of standard /o/, meaning there is a conflict perceived. This is expected to result in a rather large/early MMN, indicating a large perceptual difference. In summary, we predict an asymmetry in perception in vowel pair /ø/-/o/, since we predict a small/late MMN for when [o] is deviant, and a large/early MMN when [ø] is deviant.

Regarding labiality, predictions are less clear. In line with FUL, we assume [LABIAL] to be specified in the underlying representation of /ø/, and no specification for labiality for front vowel /e/. When [e] is the deviant sound and labial /ø/ is the standard sound, no mapping of labiality features takes place, simply because no labiality information is extracted from the signal of the deviant, which is necessary to initiate mapping. Mapping when [e] is deviant will thus not result in a mismatch and no phonological conflict will be perceived. However, FUL's predictions are less clear when the deviant is [LABIAL]: what should this feature be mapped onto? In FUL, the feature geometry suggests it would be compared to [CORONAL] and [DORSAL], but such mapping is not meaningful, since [LABIAL] is not mutually exclusive with either [CORONAL] or [DORSAL]. As a step in resolving this issue, we hypothesized a geometry that reflects this. With FUL's logic, however, we assume that nonlabial is not specified. Hence, when the deviant is labial, it cannot mismatch, because there is nothing in the representation to mismatch with. As such, a nonlabial deviant [e] does not initiate mapping and a labial deviant [ø] results in a no-mismatch. Thus, neither direction leads to phonological conflict, predicting symmetry. Accordingly, we predict similar degrees of conflict and hence similar MMNs for both directions of change.

However, the contrast between /e/ and /ø/ is crucial to perceive in Dutch. A phonological system that results in no phonological conflict in either direction of change seems inadequate in order to achieve successful discrimination of this lexical contrast. There seems to be something lacking to resolve this issue.

Possibly, the answer lies in a difference in nature between the place contrast and the labiality contrast. A separate LABIAL node indicates the feature [LABIAL] is not mutually exclusive with LINGUAL features [CORONAL] and [DORSAL]. FUL assumes a node in the geometry to be present only when a daughter feature is present – either specified or underspecified. A Dutch vowel is always either coronal or dorsal and hence always has a LINGUAL node, even when [CORONAL] is underspecified. This is different for the LABIAL node. A vowel may be nonlabial. This is not a natural class nor does it have extractable acoustic features and hence it is not an actual phonological feature. As such, nonlabiality is not represented whatsoever, and as a result nonlabial vowels have no LABIAL node. This indicates that the place contrast and the labiality contrast are different. The solution may be related to the node structure. We will revisit this in the Discussion section of the current chapter and in Chapter 5.

**Table 2.1** Hypothesized representations and predicted outcomes.

| (1) Vowel pair | (2) Contrast(s) | (3) Featural distance | (4) Standard | (5) Deviant | (6) SR of standard | (7) UR of standard | (8) SR of deviant | (9) Matching SR of deviant to UR of standard | (10) MMN properties | (11) Asymmetry |
|---|---|---|---|---|---|---|---|---|---|---|
| /ø/~ /o/ | place | 1 | /ø/ | [o] | [CORONAL] [LABIAL] | [Φ] [LABIAL] | [DORSAL] [LABIAL] | no-mismatch match | small/late | YES |
| | | | /o/ | [ø] | [DORSAL] [LABIAL] | [DORSAL] [LABIAL] | [CORONAL] [LABIAL] | mismatch match | large/early | |
| /e/~ /ø/ | labiality | 1 | /e/ | [ø] | [CORONAL] | [Φ] *empty* | [CORONAL] [LABIAL] | no-mismatch no-mismatch | small/late | NO |
| | | | /ø/ | [e] | [CORONAL] [LABIAL] | [Φ] [LABIAL] | [CORONAL] | no-mismatch | small/late | |
| /o/~ /e/ | place labiality | 2 | /o/ | [e] | [DORSAL] [LABIAL] | [DORSAL] [LABIAL] | [CORONAL] | mismatch | large/early | YES |
| | | | /e/ | [o] | [CORONAL] | [Φ] *empty* | [DORSAL] [LABIAL] | no-mismatch no-mismatch | small/late | |

A deviant vowel is represented in square brackets – a standard vowel is indicated by slashes. SR = surface representation. UR = underlying representation. [Φ] represents an underspecified feature. In the mapping procedure, the extracted features (= SR features) of the deviant (column 8) are compared to the UR features of the standard (column 7). The result of the matching procedure is reported in column 9. How this is expected to be reflected by the MMN response is given in column 10. If both directions of change for one vowel pair result in different predicted MMN properties, this is indicated as a predicted perceptual asymmetry in column 11.

Lastly, we tested a two-feature contrast with vowel pair /e/-/o/: vowels differ with respect to place as well as labiality. Based on the place contrast, we predict an asymmetry. We expect that when coronal [e] is the deviant, coronal is extracted from the signal and will mismatch with the dorsal underlying representation of /o/. Accordingly, we expect that this change will result in a large/early MMN response. When [o] is the deviant, we expect both [DORSAL] and [LABIAL] to be extracted from the signal. [DORSAL] will not mismatch with underspecified coronal place feature for /e/. Assuming labiality is not specified (empty) in UR of /e/, based on FUL's mapping procedure we predict that a [LABIAL] deviant will also not mismatch with the representation of /e/. Hence, when [o] is the deviant, place nor labiality mapping will result in a mismatch. For this direction of change we consequently predict a relatively small/late MMN response. In summary, based on phonological features and their representations in FUL, we predict an asymmetry for /e-o/ which is driven by the place contrast. However, this was not found for German (Eulitz & Lahiri, 2004). The authors suggest that this may be due to the large acoustic distance between the two vowels. The same could be the case in Dutch. As such, Dutch data also may not show a perceptual asymmetry for the two-feature contrast. Alternatively, it may be the case that asymmetries in vowel contrasts are most pronounced for a minimal, single-feature distance, as suggested by an fMRI study looking at discrimination of vowel contrast between four vowels /e/, /ø/, /o/ and /u/ in German (Scharinger, et al., 2016). Vowels differed regarding one, two or three features. In general, contrasts elicited stronger responses when the vowel changed from more specified to less specified. This pattern paralleled previous findings from Eulitz and Lahiri (2004). This effect of specificity interacted with featural distance: the increased response in one direction was only present in single-feature contrasts. For pairs with feature distances larger than one, Scharinger et al. (2016) observed that a change from less to more specific elicited stronger responses than vice versa. Thus, they conclude that with feature distances larger than one, the underspecification interpretation does not hold for their data. Also, as suggested for the labiality contrast above, our hypothesized phonological system still seems to lack an account to explain how the labiality contrast is discriminated, whilst labiality appears to play an important role in Dutch. In sum, it remains an empirical question whether the the place contrast will drive a perceptual asymmetry between /e/ and /o/.

## Method

We measured perceptual difference between three tense mid vowels /e, ø, o/ in Dutch by means of a mismatch negativity (MMN) experiment, using a passive oddball paradigm. The experiment was similar to the EEG study in German by Eulitz and Lahiri (2004). The three vowels form three vowel pairs or contrasts. Both directions of discrimination were tested for all three vowel pairs and we assessed the presence of asymmetries.

This experiment was conducted at the Donders Centre for Cognitive Neuroimaging in Nijmegen (The Netherlands). The study was approved by the local ethics committee Commissie Mensgebonden Onderzoek (CMO) Arnhem-Nijmegen, The Netherlands, under the general ethics approval (Imaging Human Cognition, CMO 2014/288), and the experiment was conducted in accordance with these guidelines.

**Subjects**

Seventeen right-handed adult native Dutch speakers (10 female, mean age = 20 y.o.; SD = 1.9; age range 18-26) were included in the final analysis. Subjects grew up speaking only Dutch at least until the age of twelve. Dialect speakers were excluded. Participants reported normal hearing, and no language or speech impediments (e.g. dyslexia, cleft palate, DLD), and subjects had never received any speech therapy. They had no background in linguistics. Prior to participating in the experiment, subjects were screened on EEG-compatibility (e.g. with respect to claustrophobia or epilepsy), and signed an informed consent. Subjects received a financial reimbursement or study course credits. Participants were recruited using the Radboud Research Participation System (SONA Systems Ltd).

Power analysis was done in G*power, based on values for the F-test for the mean amplitude of Fz reported in Eulitz and Lahiri (2004), using partial omega squared of 0.245 (based on $F(1,11) = 5.12$), see Lakens (2013). With resulting effect-size ƒ(U) of 0.57, for a power of 0.8 fourteen participants were required as a minimum for the current study. We used six versions (lists) of the experiment to counterbalance order of presentation (each testing all conditions), and we preferred to test each version with an equal number of participants. We therefore aimed for eighteen participants – three participants per list version.

In total, twenty-four subjects participated in this EEG study. Six subjects were excluded due to technical errors (4 subjects), the use of antidepressants (1 subject), and too noisy EEG data (1 subject). These subjects were replaced to get to the aimed eighteen subjects. During analysis, one final subject was excluded due to the failure to show an MMN response in any condition nor overall, resulting in seventeen included participants in the final analysis.

**Procedure and design**

Subjects participated in a passive oddball paradigm. Other than staying awake, there was no overt task. Their electrical brain activity was recorded by means of EEG while they listened to streams of vowels in isolation. Stimuli were presented using Presentation Software (version 18.2 02.18.16, www.neurobs.com). The complete EEG-recording took roughly 1,5 hours. In every block, one vowel category occurred frequently (the *standard*), interspersed with tokens of a vowel category that occurred infrequently (the *deviant*). For instance, participants heard /e/ as a standard and [o] as deviant sound in one block.

During the experiment, participants were seated in front of a computer screen (Benq XL2420Z - 24 inch) in a sound attenuated booth. Participants were instructed to sit as still and relaxed as possible, to minimize movement artifacts, eye movements and blinks. Auditory stimuli were presented binaurally through over-ear Sennheiser headphones. Participants watched a silent movie to stay engaged and awake. The film was presented at screen centre with half screen width to minimize saccades. Viewing distance was approximately 100 cm.

We tested sensitivity to the contrasts between tense mid vowels /e/, /ø/ and /o/. For each vowel pair, two conditions were tested (see Table 2.2), referring to different directions of change: e.g, for vowel pair /e,ø/, /e/ serves as a standard and [ø] as a deviant in one direction of change, whilst in the opposite direction of change /ø/ is the standard and [e] is the deviant. Having two directions of change for all three vowel pairs (or contrasts) results in six conditions. Each condition was tested in a separate block. Each participant was tested on all six conditions (within-subject). We used six different orders of blocks – counterbalanced across subjects. Two successive blocks never had the same standard vowel.

**Table 2.2** Overview of experimental conditions.

| Contrast | Condition | Standard | Deviant | Contrastive feature(s) |
|---|---|---|---|---|
| /ø/-/o/ | [ø]/o/ | /o/ | [ø] | place |
| | [o]/ø/ | /ø/ | [o] | |
| /e/-/ø/ | [e]/ø/ | /ø/ | [e] | labiality |
| | [ø]/e/ | /e/ | [ø] | |
| /e/-/o/ | [e]/o/ | /o/ | [e] | place & labiality |
| | [o]/e/ | /e/ | [o] | |

A vowel in // represents the standard sound. Vowels in square brackets [ ] represent the deviant vowel. E.g. condition [e]/o/ measures the response to vowel [e] as a deviant in a stream of /o/ stimuli.

In every block, 1000 vowels were presented in a passive oddball paradigm, with 85% standards and 15% deviants. An inter stimulus interval of 700 ms was used. Note that due to misinterpretation this ISI is longer than the 500 ms ISI in Eulitz and Lahiri (2004). We used six different orders of blocks – counterbalanced across subjects. Two successive blocks never had the same standard vowel. Each block lasted for ~15 mins. Participants were free to take a break in between the blocks. Participants themselves indicated when they were ready to start the next block. For each of the six different block orders, different pseudo-randomized within-block stimulus lists were used. Pseudo-randomized stimulus lists ensured that every block started with at least three tokens of the standard, and a deviant was always followed by at least two and maximum eleven standards before another deviant was presented. Deviants occurred unpredictably.

**Stimuli**

Our standard and deviant stimuli were three different tokens of the three Dutch vowels [e] as in *zeef* (*sieve*), [ø] as in *deuk* (*dent*) and [o] as in *poot* (*paw/leg*), spoken by a male Dutch native speaker. By using three tokens of each vowel, acoustic variability was introduced to simulate more natural speech perception conditions, and to force the processing system to map the incoming acoustic signals onto more abstract representations, rather than focus on properties unimportant in verbal processing (Eulitz & Lahiri, 2004). The vowels were recorded in isolation, and in /hV/ context. By means of a behavioural pretest we ruled out any perceptual differences due to differences in original context (see Appendix 2). For more details regarding stimulus creation, see Appendix 1.

Stimuli had a duration of 200 ms with 50 ms offset ramps. F0 was similar (~112 Hz) in all vowel categories. dBA values measured at the headphones for all vowels were within a 1 dB range of 64 dB. Since different vowel categories have different frequency characteristics, which could lead to differences in perceived loudness, we ran another behavioural pretest to rule out such differences. In addition, we controlled for degree of within-category variation by means of a pretest as well. Both these pretests are reported in Appendix 2.

The vowel categories mainly differ with respect to F2 and F3, which are related to place and labiality features. In Figure 2.3, three tokens of each of the three vowel categories are placed in F2/F3 vowel space. Figure 2.3 shows that the acoustic distance between [ø] and [e] on the one hand, and between [ø] and [o] on the other was similar. Acoustic stimulus characteristics regarding F0, intensity, as well as formant frequencies are reported in more detail in Appendix 3.

**EEG recording**

EEG data were recorded using Brain Vision Recorder (Brain Products GmbH, Munich, Germany) with 64 active electrodes (ActiCAP, equidistant – Brain Products) against left mastoid as an online reference, using a sampling rate of 500 Hz. Eye movements and blinks were tracked by 4 EOG (electrooculography) electrodes: one above and one below the left eye tracking vertical eye movement and blinks, and one left of left eye and one right of right eye tracking horizontal eye movements. Impedance levels of <20 kΩ were adopted for each channel, and we employed <5 kΩ impedance for reference electrodes on both mastoids.

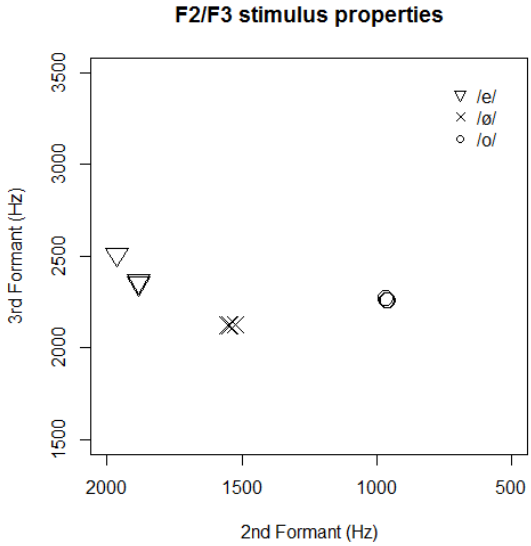



**Figure 2.3** Three tokens per vowel category in F2/F3 space.

## Data analysis

Data were analysed using Brain Electrical Source Analyses 6.0 (BESA; MEGIS Software GmbH, Gräfelfing, Germany). EEG data were re-referenced against linked mastoids. Filter and data cleaning parameters are based on analyses of the German experiment (Eulitz & Lahiri, 2004). Data were band pass filtered with low cut-off at 0.1 Hz (6 dB/oct slope) and high cut off 30 Hz (12 dB/oct slope). Epochs containing large non-eye artifacts found by visual inspection were discarded. Independent component analysis (ICA) was performed to correct for eye movement and blink artifacts. Remaining epochs with artifacts exceeding 100μV within the 100 – 650 ms time window were discarded.

Each vowel category served as a standard in two blocks, for example: /e/ occurs as a standard in a block with [o] as a deviant as well as in a block with [ø] as a deviant. Responses the first standard stimulus in a block were removed, as well as to the first standard immediately after a deviant stimulus. The event-related potential (ERP) waveforms of the same standard vowel in different blocks were not statistically different. Therefore, for standard vowels, the ERP was calculated over both blocks for every vowel. This way, we have the most robust measure for standards.

Event-related brain responses to different vowels show different waveform morphologies that may have nothing to do with any change detection process (e.g. Roberts, Ferrari, Stufflebeam & Poeppel, 2000). The mismatch negativity (MMN) response is not a regular ERP component, but is in fact a difference waveform between two ERPs (standard – deviant). To avoid effects due to the use of different vowels, the MMN is calculated by subtracting the ERP to a vowel when it is a deviant in a particular context (= the vowel that is the standard in that block), from when that same vowel serves as a standard itself – so in different blocks. For example, the MMN for the condition where [e] serves as a deviant and /o/ as standard is calculated as follows:

As such, the MMN reflects *change* detection only, and we assess the impact of the context of the deviant. We calculated MMNs for each condition for each individual participant.

For each experimental condition, we used two dependent variables for our analyses, which are similar to the ones used in Eulitz and Lahiri (2004) :

1. ***MMN latency*** measured at maximum negative amplitude at frontal electrode (Fz) in the latency range from 100-250 ms after stimulus onset, based on a window around the mean MMN latency over all conditions.
2. ***MMN amplitude*** (μV) at Fz position measured as the mean amplitude across 80 ms centred at the mean MMN latency across subjects in the corresponding experimental condition.

These two parameters were subjected to a two-way repeated measures ANOVA. The ANOVA was restricted to the two pairs of inversion with a similar acoustic change: /e,ø/ and /ø,o/. The third vowel pair /e,o/ shows markedly larger acoustic difference between the two vowels and was therefore statistically tested separately. The ANOVA had two within-subject factors:

1. ***Pair-of-Inversion*** showing a similar acoustic change: [e]/ø/ versus [ø]/e/ and [ø]/o/ versus [o]/ø/.
2. ***Direction-of-Change*** of F2 frequency between standard and deviant: ascending in [e]/ø/ and [ø]/o/, but descending in [ø]/e/ and [o]/ø/.

This ANOVA assesses whether the MMN differs between vowel pairs (*Pair of Inversion*) and whether there are general acoustic asymmetric influences on the MMN (*Direction of Change*), or different asymmetries for different vowel pairs (interaction *Pair-of-Inversion* and *Direction of Change*). Asymmetries in the MMN for the different vowel pairs subsequently were assessed by directly comparing the MMN characteristics (latency and amplitude at Fz) for the different directions of change by means of paired sample t-tests (planned comparisons).

## Results

The overall MMN showed a clear grand average with peak latency of 218 ms, with a clear frontal topography. Voltage topography maps show typical MMN topographies (e.g. Näätänen, Paavilainen, Rinne, & Alho, 2007; Näätänen, 2001) as is displayed in Figure 2.4 – with a predominant influence of left and right hemispheric temporal generators on the MMN, similar to Eulitz and Lahiri (2004). Visual inspection of grand average MMN waveforms showed a later peak latency than reported in Eulitz and Lahiri (2004), where all conditions had an MMN peak latency shorter than 170 ms. As our grand average MMN was later, we used a later time window to find peak latency values for each condition in each participant. We used a 100-250 ms window, covering the entire window of where a typical MMN peak would occur.

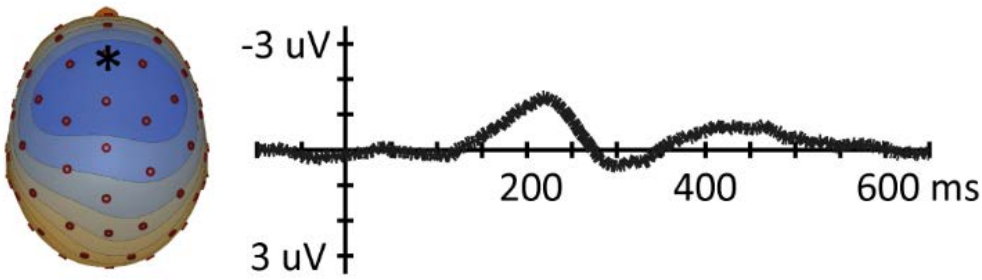



**Figure 2.4** (Left) Overall MMN voltage topography map (μV) for the 5 ms time window around the average peak latency (218 ms). Location of Fz electrode indicated with *. Blue indicates negative potentials, red positive potential. (Right): MMN waveform at frontal electrode position (Fz).

The two-way ANOVA revealed statistically significant interactions for *Pair of Inversion*Direction of Change* for MMN amplitude at the Fz ($F(1/16)$ = 26.3; p = 0.000) as well as for MMN peak latency at Fz ($F(1/16)$ = 6.3; p = 0.023). In addition, a main effect for *Pair of Inversion* (i.e. vowel pair) appeared significant for amplitude ($F(1/16)$ = 12.09; p = 0.003), but not for latency. *Direction Of Change* appeared non-significant in both amplitude and latency measures. The attested main-effect of *Pair of Inversion* shows that the two vowel pairs with similar acoustic distance behave differently. The lack of a main effect of Direction of Change implies that results cannot be explained merely on acoustic change in F2. The attested interaction in all measures implies that in processing of these vowel pairs, the impact of direction of change differs for different contrasts.

Planned comparisons of directions of change to assess asymmetries were tested with t-tests. Results of MMN amplitude and latency measured at Fz for each condition are reported in Table 2.3 below. MMN topographical plots and waveforms are presented in Figures 2.5, 2.6, and 2.7 for each contrast separately.



**Table 2.3** Table of results: MMN amplitude and peak latency at Fz for all six experimental conditions. Significance of paired sample t-tests comparing directions within one vowel pair is indicated with *.

| | Contrast | Distinctive feature | Condition | Standard | Deviant | Mean Fz Amplitude ± SEM (μV) | Mean peak latency ± SEM (ms) |
|---|---|---|---|---|---|---|---|
| 1 | /ø/-/o/ | place | [ø]/o/ | /o/ | [ø] | -1.91 ± 0.44 | 210.0 ± 6.59 |
| | | | [o]/ø/ | /ø/ | [o] | 0.28 ± 0.31 | 238.9 ± 9.09 |
| 2 | /e/-/ø/ | labiality | [e]/ø/ | /ø/ | [e] | -1.08 ± 0.31 | 216.4 ± 6.92 |
| | | | [ø]/e/ | /e/ | [ø] | -2.81 ± 0.40 | 218.9 ± 7.49 |
| 3 | /e/-/o/ | place labiality | [e]/o/ | /o/ | [e] | -1.49 ± 0.22 | 202.6 ± 7.88 |
| | | | [o]/e/ | /e/ | [o] | -2.06 ± 0.26 | 211.8 ± 5.65 |

Vowels in square brackets [ ] represent the deviant vowel to which the MMN response is measured in a certain context. A vowel in // represents the standard (=context).

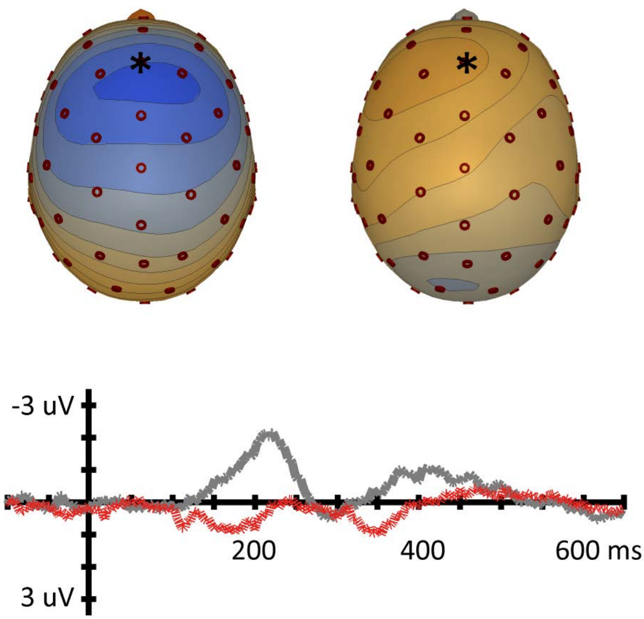



**Figure 2.5** Mismatch negativity for the place contrast (vowel pair /ø-o/). Upper panel:Topographic voltage plots of the MMN at the average MMN peak latency for both conditions left: [ø]/o/, 210 ms; right: [o]/ø/, 239 ms). Blue indicates negative potential (μV), red indicates positive potential. Electrode Fz is indicated with *. Lower panel: MMN waveforms for the Fz electrode [ø]/o/ (grey) and [o]/ø/ (red).

For the place contrast /ø-o/, a paired t-test showed a significantly larger MMN amplitude at Fz for condition [ø]/o/ (M = -1.19 μV; SD = 1.81) than for condition [o]/ø/ (M = 0.28 μV; SD = 1.31); t(16) = 4.281, p < 0.001. The peak latency measure resulted in a significant asymmetry in the same direction: a shorter latency was found for [ø]/o/ (M = 209.9 μV; SD = 27.2) than for [o]/ø/ (M = 238.9 μV; SD = 37.5); t(16) = 2.823, p = 0.012.

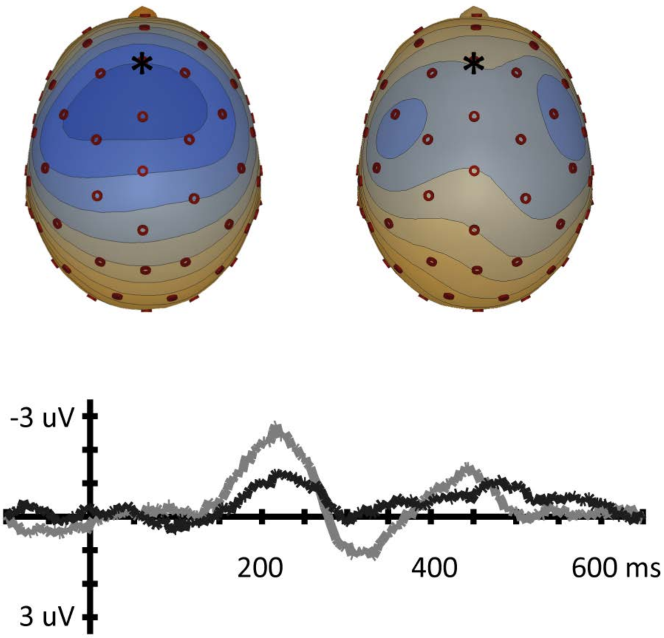



**Figure 2.6** Mismatch negativity for the labiality contrast (vowel pair /e-ø/). Upper panel:Topographic voltage plots of the MMN at the average MMN peak latency for both conditions left: [ø]/e/, 219 ms; right: [e]/ø/, 216 ms). Blue indicates negative potential (μV), red indicates positive potential. Electrode Fz is indicated with *. Lower panel: MMN waveforms for the Fz electrode [ø]/e/ (grey) and [e]/ø/ (red).

For the labiality contrast /e-ø/, a paired t-test on Fz amplitude revealed a significantly larger MMN response for [ø]/e/ (M = -2,18 μV; SD = 1.66) compared to condition [e]/ø/ (M = -1.08 μV; SD = 1.28); t(16) = -4,053, p < 0.001. Thus, the results show a significant asymmetry between [ø]/e/ and [e]/ø/ with respect to amplitude. Fz MMN peak latency showed no significant difference between [ø]/e/ (M = 218.9 μV; SD = 30.9) compared to [e]/ø/ (M = 216.4 μV; SD = 28.5); t(16) = 0.284, p = 0.78.

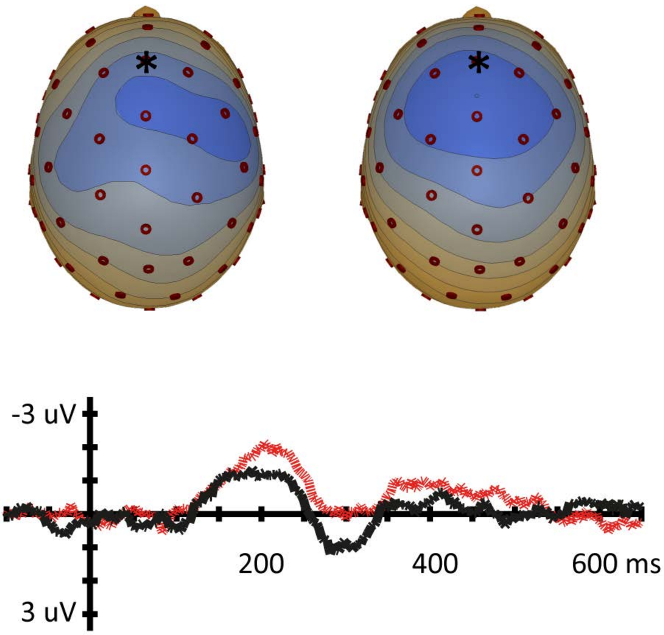



**Figure 2.7** Mismatch negativity for the two-feature contrast (vowel pair /e-o/). Upper panel: Topographic voltage plots of the MMN at the average MMN peak latency for both conditions left: [e]/o/, 203 ms; right: [o]/e/, 203 ms). Blue indicates negative potential (μV), red indicates positive potential. Electrode Fz is indicated with *. Lower panel: MMN waveforms for the Fz electrode [e]/o/ (grey) and [o]/e/ (red).

For two-feature contrast /e-o/, the paired t-test did not show significant results: Fz amplitude for [e]/o/ (M = -1.49 μV; SD = 0.93) was not significantly different from [o]/e/ (M = -2.06 μV; SD = 1.1); t(16) = -1.857, p = 0.083. The t-test on Fz latency measure for [e]/o/ (M = 202.6 μV; SD = 32.5) and from [o]/e/(M = 211.8 μV; SD = 23.3) was not significant either; t(16) = 1.051, p = 0.309.

To conclude, we found significant asymmetries for both single-feature contrasts, but no significant asymmetry was found for the two-feature contrast.

## Discussion

The present study set out to test how place and labiality contrasts are represented and perceived by Dutch listeners. We conducted an EEG study similar to what has previously been done in German by Eulitz and Lahiri. On the surface, German and Dutch vowel systems appear similar; both have a three-way contrast between /e, ø, o/. Yet, looking at the linguistic system as a whole, the nature of the front rounded vowels appears different in Dutch and German. Hence, we were interested to see whether these differences are reflected in processing, in particular when it comes to the labiality contrast. Hypothesized representations and predicted mapping outcomes are recapitulated in Table 2.4.

**Table 2.4** Hypothesized representations and predicted mappings for the three vowel contrasts in two directions of change.

| | | **(1) Place: /o-ø/** | | **(2) Labiality: /e-ø/** | | **(3)Two-feature: /e-o/** | |
|---|---|---|---|---|---|---|---|
| **UR** | PLACE | [DORSAL] | [Φ] | [Φ] | | [DORSAL] | [Φ] |
| | LABIALITY | [LABIAL] | [LABIAL] | *empty* | [LABIAL] | [LABIAL] | *empty* |
| | standard sound | /o/ | /ø/ | /e/ | /ø/ | /o/ | /e/ |
| | **Mapping** | ‡ | ↿ | ↿ | ↿ | ‡ | ↿ |
| **SR** | deviant sound | [ø] | [o] | [ø] | [e] | [e] | [o] |
| | LABIALITY | [LABIAL] | [LABIAL] | [LABIAL] | *nonlabial* | *nonlabial* | [LABIAL] |
| | PLACE | [CORONAL] | [DORSAL] | [CORONAL] | | [CORONAL] | [DORSAL] |
| Asymmetry predicted? | | YES | | NO | | YES | |

UR = Underlying Representation, features as stored in the mental representation. SR = surface representation, i.e. features as extracted from input. Vowel in // serves as standard, vowel in [] serves as deviant. Φ indicates an underspecified (default) feature. Nonlabial is included for sake of clarity, and is indicated in grey italics, because it is not assumed an extracted feature in SR. It is also assumed not specified in UR, indicated by 'empty' in grey italics. Mapping is indicated by symbols ‡ for mismatch (conflict – access to UR is blocked) and ↿ for no-mismatch/match (no conflict – access to UR provided).

An overview of the predicted MMN properties along with the actual experimental results is given in Table 2.5, with vowels serving as standards indicated in //, and a vowel serving as deviants indicated in [].

**Table 2.5** Overview of predictions and results of MMN amplitude and latency. Relative results are given, e.g. small/large and early/late.

| Vowel contrast | Contrastive features | Standard | Deviant | Prediction | | Results | |
|---|---|---|---|---|---|---|---|
| | | | | MMN amplitude and latency | asymmetry | MMN amplitude and latency | (a)symmetry |
| /ø/~ /o/ | Place | /ø/ | [o] | small amplitude<br>late peak | enhanced MMN for [ø] /o/ | (very) small amplitude<br>late peak | enhanced MMN for [ø] /o/ |
| | | /o/ | [ø] | large amplitude<br>early peak | | large amplitude<br>early peak | |
| /e/~ /ø/ | labiality | /e/ | [ø] | small amplitude<br>late peak | similar MMN in both directions of change | (very) large amplitude<br>n.a. | enhanced MMN for [ø] /e/ |
| | | /ø/ | [e] | small amplitude<br>late peak | | small amplitude<br>n.a. | |
| /o/~ /e/ | place & labiality | /o/ | [e] | large amplitude<br>early peak | enhanced MMN for [e] /o/ | large amplitude<br>n.a. | No asymmetry. Trend of enhanced MMN for [o] /e/ |
| | | /e/ | [o] | small amplitude<br>late peak | | (very) large amplitude<br>n.a. | |

N.a. = not applicable is noted in case there is no asymmetry for latency in that vowel pair. The term enhanced MMN refers to larger and/or earlier MMN peaks.

We found significant asymmetries for both single-feature contrasts (place contrast /ø-o/, and labiality contrast /e-ø/). No significant asymmetry was found for the two-feature contrast (/e-o/). Regarding the language universal contrast of place, the asymmetry in vowel pair /ø-o/ meets our expectations. Results diverge from our predictions for the labiality contrast /e-ø/ and the two-feature contrast /e-o/. In the next section we propose an account explaining current outcomes.

**Explaining asymmetries**

For the **place contrast** between /o/ and /ø/ we replicated findings in German by Eulitz and Lahiri (2004); a change from standard /o/ to deviant [ø] shows a larger MMN amplitude and shorter peak latency compared to the reverse [o]/ø/. One direction of change yields a phonological conflict. This provides evidence for [CORONAL] underspecification in Dutch, similar to what has been argued for by Eulitz and Lahiri (2004) for German vowels. A perceived [DORSAL] feature yields a no-mismatch with an underspecified [CORONAL] feature, whereas a perceived [CORONAL] feature in the surface representation yields a mismatch with a specified [DORSAL] feature in de underlying representation (see Figure 2.8). A similar finding has been reported by De Jonge and Boersma (2015) for French, who attested an asymmetry between coronal /y/ and dorsal /u/: they also report an enhanced MMN when the deviant was coronal.

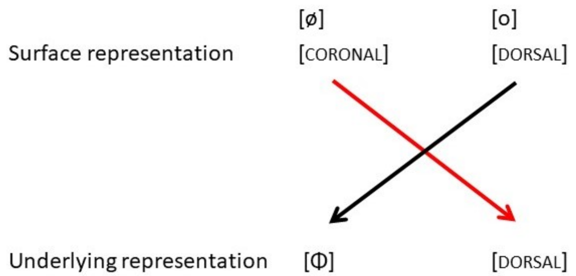



**Figure 2.8** Asymmetrical perception of the vowel pair /ø/-/o/. Black lines indicate no-mismatch, the red line indicates a mismatch. [ø] indicates underspecification.

Regarding the language universal contrast of place, the asymmetry in vowel pair /ø-o/ supports the notion of coronal underspecification as assumed by the FUL model, and the results also fit with the assumption that [DORSAL] is specified. The findings support the notion of universal underspecification of coronal.

Results diverge from our predictions for the **labiality contrast /**e-ø/. We found an asymmetry for the labiality contrast; There was an enhanced MMN when deviant [ø] is mapped onto the representation of standard /e/. This enhanced MMN indicates a mismatch/conflict when the deviant is labial. However, nonlabial /e/ has no labiality information stored in UR, so there is no feature present for SR feature [LABIAL] to mismatch with. Thus, we did not predict conflict in this direction of change.

However, we already noted that it is puzzling how listeners perceive the labiality contrast if no direction of change leads to conflict in the labiality contrast. Data now show that indeed at least one direction of change does lead to phonological conflict. This supports our assumption that the labiality contrast plays a crucial role in Dutch vowels, and perceiving this contrast at a phonological level is required (we will come back to how this may work below). This asymmetry was not reported for German. Rather, Eulitz and Lahiri (2004) argue that this is a no conflict situation. As such, German and Dutch processing of the labiality contrast seems to differ. This supports our view that labiality plays a different role in both languages.

For the **two-feature contrast** /e-o/ a similar asymmetry was predicted as for single-feature contrast /ø-o/ (place): a coronal-dorsal asymmetry. However, results do not show this. Rather, there is a trend in opposite direction: a larger MMN response when deviant is [DORSAL, LABIAL] than when the deviant is [CORONAL]. Lack of significance could be due to the large acoustic distance in this contrast, similar to the explanation Eulitz and Lahiri provide for the lack of significance for the /e-o/ contrast in German or due to a two-feature difference as proposed by Scharinger et al. (2016). Unlike Eulitz and Lahiri, we found a trend in opposite direction from a coronal-dorsal asymmetry. Instead, it is in similar direction as the labiality asymmetry between /e/ and /ø/. We will argue that results imply an important role of the labiality contrast explaining the outcomes for both the labiality contrast /e-ø/ and the two-feature contrast /e-o/.

Overall, labial and coronal features are not regarded as mutually exclusive. Else, coronal SR features would have caused a conflict with labial UR features in the contrast between /e/ and /ø/, resulting in an enhanced MMN when the deviant was /e/. We therefore argue that the findings are in line with the hypothesized geometry, assuming coronal and dorsal to be separated from labial. This assumption is shared by various researchers (e.g. Browman & Goldstein, 1989; Keyser & Stevens, 1994). Following this hierarchy, we propose the following representations for Dutch vowels /e/, /ø/, and /o/.

Surface representations:

```
     /e/                /ø/                 /o/
      |                  |                   |
     ART                ART                 ART
      |                 / \                 / \
   LINGUAL        LABIAL   LINGUAL    LABIAL   LINGUAL
      |              |        |          |        |
  [CORONAL]      [LABIAL] [CORONAL]  [LABIAL]  [DORSAL]
```

Underlying representations ([Φ] indicates an underspecified feature):

```
     /e/                /ø/                 /o/
      |                  |                   |
     ART                ART                 ART
      |                 / \                 / \
   LINGUAL        LABIAL   LINGUAL    LABIAL   LINGUAL
      |              |        |          |        |
     [Φ]          [LABIAL]   [Φ]     [LABIAL]  [DORSAL]
```

Note that we follow FUL's assumption that nonlabial is not an existing feature and for nonlabial vowels, this property is not specified. Hence, for nonlabial vowels, no LABIAL node is part of the representation – see SR and UR representations of /e/.

However, this does not yet adequately explain the findings for the labiality contrast. How does conflict arise when the standard nonlabial vowel is followed by a [LABIAL] deviant?

We hypothesized the mapping procedure assumed in the FUL model. Features extracted from the acoustics are compared to features stored in the underlying representation. The goal of this feature mapping is to deselect unwanted candidates, and to limit the number of word candidates for word recognition. Of course, only the relevant comparisons are made. After all, when for instance a feature [CORONAL] is extracted, only mapping to a mutually exclusive underlying feature like [DORSAL] would result in a meaningful/informative mismatch, whereas mapping [CORONAL] onto a feature like [HIGH] is neither efficient nor informative. In order for the mapping process to result in meaningful (no-)(mis)matches, the scope of the comparison of surface features must be defined. In general, FUL assumes features underneath a single node to be mutually exclusive. As such, nodes can define which feature mappings may be meaningful. With our hypothesized feature hierarchy, nodes indeed indicate which features are mutually exclusive for Dutch vowels.

Our geometry does not yield different predictions for the place contrast than FUL would, as [DORSAL] and [CORONAL] still share a node. However, an important consequence of the geometry occurs with respect to labiality, as [LABIAL] is the only feature underneath the LABIAL node. As previously mentioned, FUL only assumes a node present when a daughter feature is present. Nonlabial is not considered a phonological feature and is thus not present in any level of representation. As a result, a vowel is either [LABIAL] with a LABIAL node, or is nonlabial and hence without a LABIAL node. This is a crucial difference with the place contrast: the LINGUAL node is always present. Thus, we can consider place and labiality different types of contrasts that may require different ways of processing to be perceived.

Since nodes indicate which phonological features are mutually exclusive, the geometry provides information regarding what should be mapped to what. Coronal or dorsal SR features should be mapped onto features below the LINGUAL node in order to perform meaningful mapping. Since vowels are always (underspecified) coronal or dorsal in Dutch, the LINGUAL node is always present. But if the nodes inform us about what is mutually exclusive, absence of the LABIAL node in nonlabial vowels may be problematic. And this may just hold the key to how the contrast is perceived. If we assume that SR features are supposed to be mapped onto features below their own parental node, the following would happen: When the acoustic feature corresponding to [LABIAL] is extracted from the signal and hence becomes part of the surface representation, mapping it to an underlying representation with [LABIAL] results in a match. When the underlying representation however lacks a LABIAL node (i.e. when UR is nonlabial), mapping cannot take place. This implies a phonological discrepancy between the surface and underlying representation, and we argue that such a case also indicates a phonological conflict in perception (see Figure 2.9)

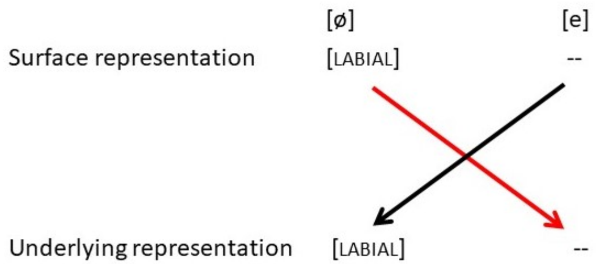



**Figure 2.9** Asymmetrical perception of the vowel pair /ø/-/e/. Black lines indicate no-mismatch, the red line indicates a mismatch. –– indicates no feature selected (SR) and node lacking (UR).

Following this logic, results for the two-feature contrast can be explained as well. We found no significant asymmetry for this two-feature contrast. From the discussion of the two vowel pairs with single-feature contrast, this may not come as a surprise. For the place contrast, we would expect an enhanced MMN in the condition /o/[e], but not vice versa. However, based on labiality, we would expect an enhanced MMN in the condition /e/[o], as shown in Figure 2.10. Mapping of the labial SR feature of [o] cannot take place because /e/ does not have a LABIAL node in UR. An explanation for the lack of a significant asymmetry may thus be that both contrasts cancel each other out; /o/[e] results in a mismatch based on coronal/dorsal; whereas for /e/[o] the mapping procedure is aborted based on labiality differences. As such, both directions of change are phonologically conflicting. In accordance with this, we found large MMNs for both directions of change for the two-feature contrast.

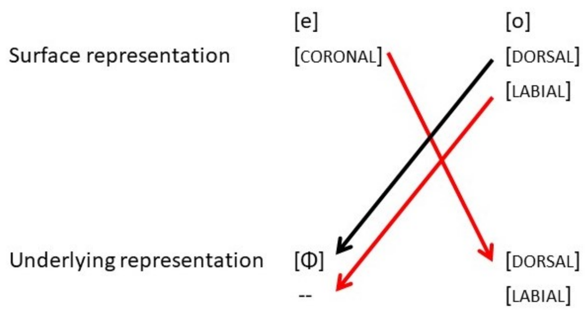



**Figure 2.10** Asymmetrical perception of the vowel pair /ø/-/e/. Black lines indicate no-mismatch, the red line indicates a mismatch. –– indicates no feature selected (SR) and node lacking (UR). [Φ] indicates underspecification. The matching for place of articulation and labiality both show phonologically conflicting results.

If aborted mapping is interpreted as a larger conflict than a mismatch (possibly because the stage of mapping is not reached as the discrepancy is already at the level of the feature geometry/nodes) perception could be asymmetrical; a larger MMN for a conflict due to aborted mapping than for a conflict due to mismatch. We found a larger MMN response when the deviant was labial than when the deviant was not labial. This pattern was found for both the single-feature labiality contrast (/e-ø/) and the two-feature contrast (/e-o/), but for the latter the asymmetry was not significant. Since in the two-feature contrast both directions of change are conflicting and both directions result in rather large MMNs, it is not evident that an asymmetry would be statistically revealed. For a slight difference in weighting, a lot of data and/or sensitive measure would be required. Since we found no significant asymmetry in the two-feature contrast, we did not find evidence for a difference in weighting between a 'regular' mismatch and aborted mapping. We potentially lacked the power to find such a subtle difference between these two types of conflict.

**Alternative explanations**

The literature mentions at least two other types of explanations for attested asymmetries in vowel perception: phonetic saliency and frequency of use (see also Chapter 1). Phonetic saliency or ease of discrimination has been used to account for asymmetries in vowel perception, among others by proponents of the Natural Referent Vowel (NRV) framework (Polka & Bohn, 2011; Masapollo, Polka, & Molnar, 2017; Masapollo, Polka, & Ménard, 2017) (see also Chapter 1). It assumes that vowels with formant frequencies closer together have focalized energy, and hence are more salient in perception than vowels with formants further apart. Consequently, changes from less to more focal vowels are easier to discriminate. In light of the current study, a larger MMN is expected when the standard is non-focal, and the deviant is focal. Regarding the place contrast: based on F1-F2, /o/ is more focal than /ø/. Based on F2-F3 /ø/ is more focal than /o/. Regarding the labiality contrast: Based on the convergence or closeness of F2 and F3, /e/ is more focal than /ø/. Based on the F1-F2 dimension, /ø/ is more focal than /e/, but the formants do not get close, and it remains the question whether the formants lead to focalized energy. In other words, predictions for both the place and the labiality contrast are not straight forward. Current findings for the labiality contrast between /e/ and /ø/ would be in line with predictions based on F1-F2, whereas findings for the place contrast between /ø/ and /o/ would be in line with predictions based on F2-F3. It must be noted that the results reported in Polka and Bohn (2011) do not always conform to their own predictions. Notably, they report better discrimination in the direction from/ e/ to /ø/ in Danish infants, which is in similar direction as our labiality asymmetry.

For frequency, or experienced-based theories, like for instance the Native Language Magnet (NLM) theory (Kuhl, 1991; Kuhl et al., 2008) it is usually assumed that category building is based on distributions in the input, and more frequent vowels have a stronger magnetic effect, warping perceptual space around them. Therefore, poorer discrimination is expected in the direction from more frequent to less frequent. However, this discrimination effect usually holds for within-category discrimination. If both are categories in the language, predictions are less clear. However, one could hypothesize that frequent vowels allow for more variation, and hence it is expected that a frequent standard and an infrequent deviant would be more difficult to discriminate, and hence show a smaller MMN effect, than vice versa. Front round vowels are relatively infrequent in Dutch compared to front (unround) vowels and back (round) vowels. Relative frequencies reported by Baayen et al. (1995) illustrate that /e/ and /o/ occur roughly twice as often as /ø/ in Dutch. For both the place contrast and the labiality contrast, it would be predicted that MMN is enhanced when the deviant is not /ø/. These predictions are opposite to the findings in our experiment.

To summarize, the current NRV framework only partly predicts the attested asymmetries. On the other hand, the frequency account based on NLM does not make the right predictions for either of the single-feature contrasts.

The proposed geometry and mapping algorithm shows how listeners might evaluate the incoming signal based on its phonology, which is conform the attested asymmetries.

**Dutch versus German**

A question raised by the results is why Dutch and German listeners show this particular different perceptual behaviour. Despite great overlap in vowel inventory between German and Dutch, there are striking differences between the results in German as reported by Eulitz and Lahiri (2004) and the current findings for Dutch. Eulitz and Lahiri (2004) reported symmetrical perception for the vowel pair /e/-/ø/, which share place of articulation, and hence do not constitute a conflicting situation for place of articulation. This result was replicated for German listeners in an MMN passive oddball paradigm using both words and nonwords (Cornell et al., 2011), who similarly report symmetrical results in this condition. This contrasts with the results in the current study, which show an asymmetry for labiality.

One hypothesis is that the status of /ø/ as an underlying phoneme is different in Dutch and German. German has much fewer monomorphemic words with /ø/ than Dutch. In German, [ø] often arises as the result of morphological umlaut as in the plural in *Vogel* [o] – *Vögel* [ø] ('bird – birds'), and in the comparative form of the adjective in *hoh* [o] – *höher* [ø] ('high – higher'). Hence, [ø] often occurs in derived environments, rather than in the lexicon. Although /ø/ occasionally also occurs in nonderived (lexical) environment, its contrastive value in German is limited, in comparison to Dutch, which may explain why German and Dutch listeners react differently to [LABIAL] in the context of [CORONAL] vowels: in German it is not a strong lexical contrast, while in Dutch it is.

Another difference between the German and Dutch vowels /e, ø, o/ is that the Dutch realization of vowels can be considered semi-diphthongized (Adank et al., 2014), whilst the German ones have more stable formant frequencies. Diphthongization strengthens place features towards the end of vowels, as front vowels become even more front, impacting the second formant (F2), which is the acoustic cue for the front-back dimension (see Figure 2.11). Furthermore, vowels may become higher towards the end, acoustically reflected by a lower first formant (F1). Although formant transitions may impact the acoustic MMN, this does not explain the presence of asymmetries. Phonologically, however, these tense mid vowels are not diphthongs in Dutch; the formant transition is not mandatory.

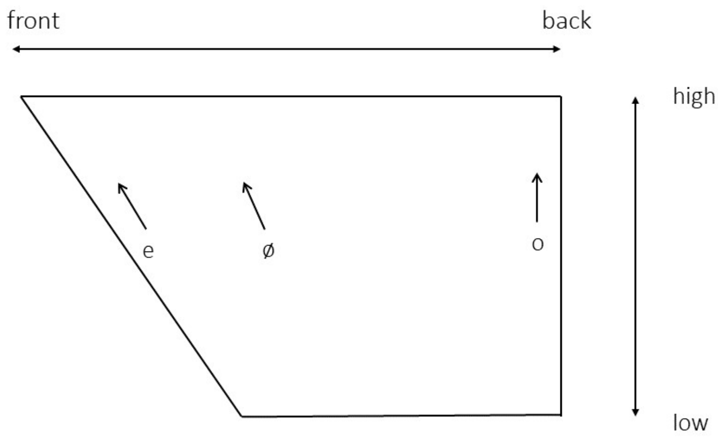



**Figure 2.11** Diphthongization of Dutch tense mid vowels in vowel space. Position of the tongue is indicated with the front-back dimension (x-axis) for place and open-close dimension (y-axis) for height. Arrows indicate shift of formant 1 (height) and formant 2 (place).

In the diagram it is apparent that place features coronal and dorsal remain stable during the whole vowel even when it is produced semi-diphthongized. The same holds for lip rounding: labial vowels are still labial along the whole vowel duration. As such, based on place of articulation features there is no reason to assume semi-diphthongization to result in different place or labiality features and hence different patterns of results for German and Dutch.

The only dimension that may change under the influence of diphthongization is vowel height. When examining formant frequencies of our stimuli (included in Appendix 3) in more detail, /ø/ could be considered to be (produced as) a high vowel since it has a low F1. Also /o/ becomes high towards the end. /e/ on the other hand is clearly not high for the entire duration of the vowel. French data supports [MID] to be underspecified for vowels, as De Jonge and Boersma (2015) report an asymmetry between high and mid vowels: the MMN is enhanced when the deviant is mid. FUL assumes mid to be either underspecified or not specified at all. This may vary across languages depending on the language's phonological system. If we assume [MID] to be underspecified, this would predict that /e/ would accept both [ø] and [o] deviants. On the other hand, a mismatch would occur when [e] serves as a deviant, as [MID] would be extracted, and would mismatch with a [HIGH] feature of /ø/ or /o/. This is not in line with the enhanced MMN for labial [ø] as deviant. If we assume mid to be not a feature (i.e. not specified), under the traditional FUL model, predictions would be the similar, since UR of /e/ has no height feature to mismatch with. So regarding height, deviants [ø] and [o] would be accepted as a variant of /e/ - also when they are [HIGH] in SR - which is not in line with current results.

However, when we assume that mapping is aborted and interpreted as phonological conflict when a node is present in SR, but not in UR, this would imply that [ø] and [o] would conflict with /e/ based on height. This would be in line with our results.

Since /ø/ and /o/ are both not obligatorily diphthongized, we would assume [HIGH] to not be stored in UR for these two vowels, and mid deviant [e] would not mismatch with either of them even if [MID] would be extracted from SR. If rather than being underspecified /e/ has no height feature in UR, this would imply that in SR nothing is extracted and no height is mapped. Thus, with respect to height features, [e] would be accepted as a variant for /ø/ and /o/ either way. However, data indicate that [e] conflicts /o/ as well as /ø/. This cannot be explained based on height. Although this conflict may be explained based on place features in case of the /e-o/ contrast, this is not possible for /e-ø/.

In sum, we cannot explain our results based on diphthongization impacting height features. However, the role of diphthongization will be revisited in Chapter 3.

**Limitations and suggestions for further research**

FUL assumes the same language universal feature tree for vowels and consonants. How does the adapted feature tree relate to consonants? Although Lahiri (e.g. 2018) states that the features [LABIAL], [CORONAL] and [DORSAL] are mutually exclusive for consonants, there are some cases that indicate otherwise: for example Cantonese, where labialized velars (dorsals) or alveolars (coronals) exist: /t/ versus /$t^w$/ versus /k/ versus /$k^w$/: a four-way distinction is made. The statement that labial is mutually exclusive with dorsal and coronal in consonants may not be true for all languages. Our proposed geometry may therefore also hold for consonants. However, we have only investigated vowels in the current study. Future research could target consonant perception in general, but in particular in languages where the ARTICULATOR features are not all mutually exclusive to see whether the adapted feature tree holds for these cases. If it does not, it should be considered whether this is because the consonants are clusters rather than single phonemes.

It could be that for contrasts similar to the labiality contrast, a similar structure is present in the feature tree to resolve the issue of how the difference can be detected perceptually when one of the contrasting features does not exist as a phonological feature in the model. A separate node that can be present or absent would resolve this; similar to the coronal-dorsal contrast one direction of change then results in a phonological conflict. An asymmetry similar to our labiality asymmetry has been reported regarding the feature [VOICE] by Van der Feest and Fikkert (2015) in Dutch toddlers' perception. They report that a change from a voiceless towards a voiced speech sound (e.g. *[b]oes instead of [p]oes (*cat*)) is perceptually more salient than an opposite mispronunciation (i.e. changing a voiced stop into a voiceless stop). This resembles the case of labiality. A change from a non-feature (voiceless) towards a specified feature [VOICE] appears to result in phonological conflict, but not vice versa. Lahiri (2018) also points out that [VOICE] could have its own node, similar to what we propose for [LABIAL]. This line of reasoning is also compatible with the perceptual asymmetries presented in Hestvik and Durvasula (2016) for English, albeit the specified feature here is [SPREAD GLOTTIS] for /t/, while /d/ is underspecified (lacking laryngeal specification). They noticed a larger MMN for [t]/d/ than for [d]/t/, although they presented a somewhat different analysis. Using a similar mapping rule, this would explain the laryngeal asymmetry (where either [VOICE] is specified as in Dutch, or [SPREAD GLOTTIS] as in English), because mapping cannot be completed when a voiced sound is heard but mapped onto a voiceless underlying representation where the VOICING node is absent, hence resulting in a phonological conflict. Whether this idea holds true more generally should be assessed in future research.

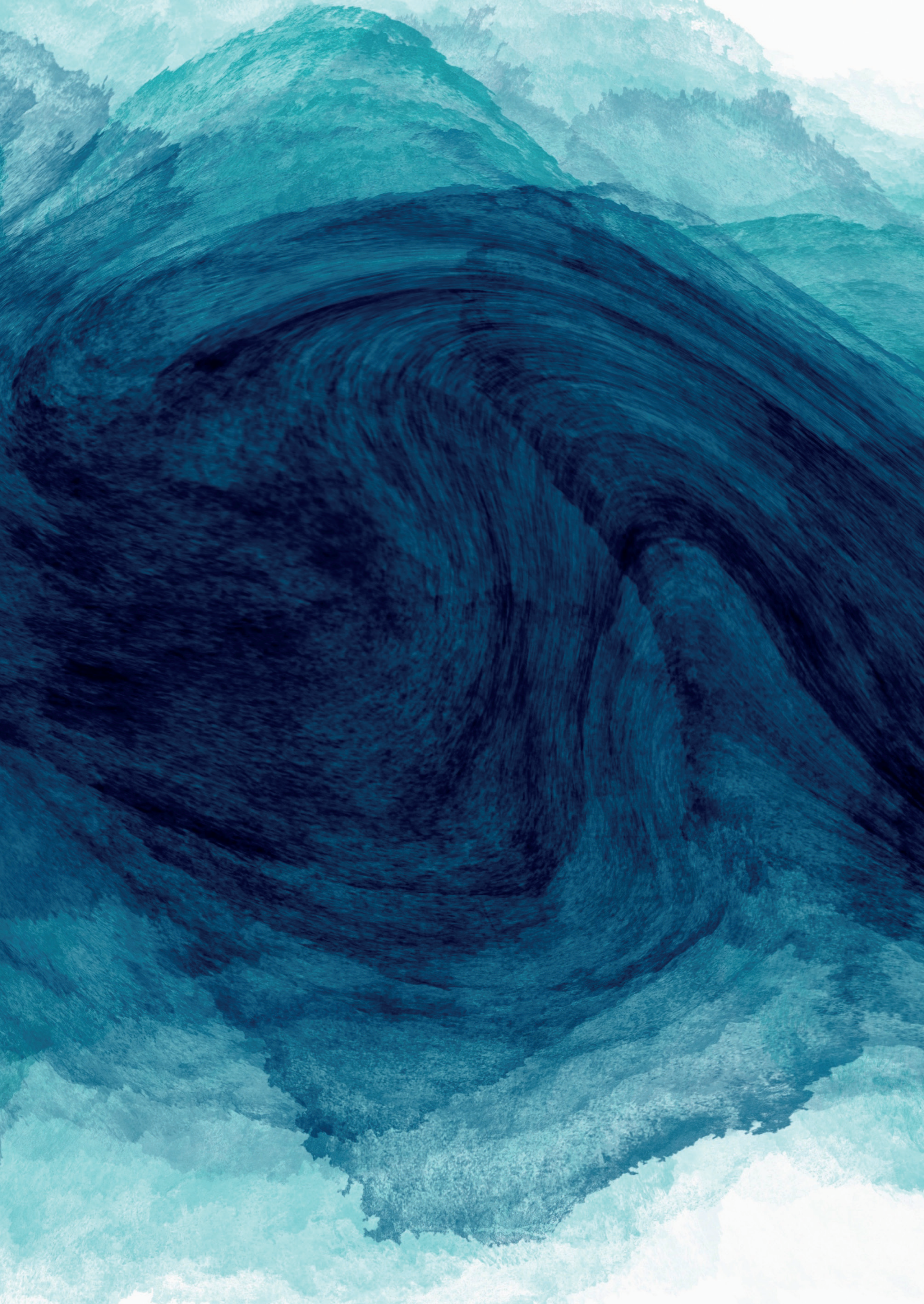
Another remaining question concerns the underlying representation of /o/. We cannot draw conclusions based on the current study on whether or not the feature labial is actually represented in UR for dorsal vowels such as /o/. As dorsal vowels in Dutch are always labial, it may be that labial is underspecified for dorsal vowels, which would be in line with FUL's assumption that redundant features may be underspecified. We tested the contrast between /e/ and /o/, where there is a contrast in labiality, but this contrast cannot resolve this question. Whether or not [LABIAL] is specified in UR for /o/ would not have changed the results/predictions, because nothing is extracted from [e] that could mismatch with [LABIAL] UR of /o/. Redundancy with dorsal would imply labial could be underspecified, but because labial is generally not default and nonlabial cannot be stored, [LABIAL] may in fact be stored for /o/. This matter is revisited in Chapter 5, also discussing a language acquisition perspective.

**Conclusion**

The current study showed evidence for asymmetrical processing of vowels, and the attested asymmetries provide further support for FUL's underspecification account. We replicated the place asymmetries in German listeners as reported in Eulitz and Lahiri (2004) for Dutch. While they did not find a labiality asymmetry (between /e/-/ø/), this asymmetry was shown in the current study. This implies a difference between German in Dutch, particularly regarding the role of labiality.

The results support our hypothesized geometry, separating LABIAL from LINGUAL features. This has at least two benefits: features under the same node are mutually exclusive. Moreover, as [LABIAL] is the only feature under the LABIAL node, its mapping must match, as a no-mismatch is not a viable option. If the perceived [LABIAL] in the surface representation cannot be matched since there is no LABIAL node in UR, the mapping is aborted, leading to a conflict, and hence a perceptual asymmetry. This is the case when the standard is /e/ (no labial node), and the deviant is [ø], with a surface feature [LABIAL]. A consequence of our interpretation is that not only a mapping mismatch (as traditionally discussed in FUL), but also an aborted mapping implies a phonological conflict.

We would like to stress that a clear three-way distinction between the terminology of *specified*, *underspecified* and *not specified* is vital. The terms *underspecified* and *not specified* have been used inconsistently or synonymously in the literature. Whether something contrastive is *underspecified* (like CORONAL) or *not specified* (like *nonlabial* or *voiceless*) has important consequences for the presence or absence of nodes in UR, and hence for predicting (asymmetries in) perception of contrasts.



# Chapter 3

## A cross-modal semantic priming study

## Introduction

In the previous chapter, we conducted an electroencephalography (EEG) experiment assessing electrical brain activity in response to vowel changes to gain insight in the featural representation of vowels in Dutch. We aimed to understand how labiality and place features are represented in Dutch by means of assessing perceptual asymmetries. The study was a replication of a study previously conducted in German listeners by Eulitz and Lahiri (2004), measuring the mismatch negativity (MMN) response as a proxy of perceptual difference between phonemes. Similar to the German listeners, our Dutch listeners showed a coronal-dorsal asymmetry attested in the contrast between coronal /ø/ and dorsal /o/; a change from a dorsal into a coronal vowel was perceptually larger (resulting in an enhanced MMN) than vice versa. However, German and Dutch diverge regarding labiality. The contrast between nonlabial /e/ and labial /ø/ resulted in an asymmetry in Dutch listeners, but not in German listeners, as reported in the study by Eulitz and Lahiri (2004).

In the current chapter, we aim to investigate whether our results from Chapter 2 are robust, by replicating the results using a different method. In particular, while Chapter 2 tested perception of vowels in isolation measuring an automatic brain response, this chapter will test vowel discrimination in words, measuring reaction times (i.e. overt response) in a lexical decision task with semantic priming. We predict the current study to show similar perceptual patterns as found in the EEG study in Chapter 2.

The findings in Chapter 2 for Dutch differed from the results in German. One aspect in which Dutch and German stimuli differ is in the diphthongization of long vowels in Dutch, whilst the corresponding vowels in German are truly monophthongs. Therefore, this study additionally aims to compare results for tense and lax mid vowels, as lax vowels are monophthongs constituting similar contrasts as the tense vowels. In short, this chapter has the following aims:

(1) We will investigate the same phonological contrasts as in Chapter 2 by means of an entirely different method, measuring overt behaviour in a semantic priming experiment rather than electrical brain activity. We will test vowel discrimination in words rather than in isolation, assessing whether our findings generalize to lexical processing;
(2) We will examine not only Dutch tense mid vowels /e, ø, o/, but also their lax counterparts /ɛ, ʏ, ɔ/, to investigate whether findings generalize to different (non-diphthongized) vowels constituting similar phonological contrasts.

**Cross-modal semantic priming**

As an alternative to using EEG to investigate perception of phonological contrasts, mispronunciation paradigms have been successful in both adults and children (e.g. Bailey & Plunkett, 2002; Van der Feest & Fikkert, 2015; Altvater-Mackensen, Van der Feest & Fikkert, 2014; Bölte & Coenen, 2000; Mani & Plunkett, 2010; Kotzor, Wetterlin, Roberts, & Lahiri, 2015; Scharinger & Lahiri, 2010). In adults, previous research on the impact of mispronunciations on lexical access has made use of semantic priming to investigate underlying phonological representations. For example, Kotzor et al. (2015) investigated duration in Bengali, whilst Scharinger and Lahiri (2010) investigated vowel height in English dialects, and Lahiri and Reetz (2002) investigated place contrasts in German word-final consonants (Lahiri & Reetz, 2002).

In cross-modal semantic priming paradigms, a subject hears a word (prime), followed by reading a written word (target). The subject performs a lexical decision task with respect to the written target word. When the prime is semantically related to the target word, this often yields priming effects; RTs are shorter than when the prime is unrelated (e.g. Scharinger & Lahiri, 2010; Kotzor et al., 2015). The line of reasoning is as follows: When the prime is heard, its phonological representation is accessed, which provides access to the lexeme and its semantics. Semantic activation spreads to semantically closely related words. Hence, the synonymous target word is pre-activated, facilitating access to the target, resulting in quicker word recognition and quicker responses. Consequently, using mispronounced primes can provide insight into whether or not a certain mispronunciation hinders word recognition: if the canonical word is not accessed in the lexicon, it will not result in semantic priming of the target word, whereas acceptable variation will still activate the canonical word and hence also prime the related word. Within one phonological contrast, one direction of change may hinder word recognition whereas the opposite mispronunciation may not. As such, cross-modal semantic priming can be used in order to investigate phonological representations and similar to the experiment in Chapter 2 it may show perceptual asymmetries.

**Current study**

In the current study, we assess adult processing of Dutch tense vowels /e, ø, o/, and their lax counterparts /ɛ,ʏ,ɔ/, by means of a cross-modal semantic priming experiment with mispronunciations, assuming similar contrasts in both sets of vowels. The three central research questions are:

RQ1. How is the vowel contrast between CORONAL and DORSAL (i.e. place) represented?

RQ2. How is the vowel contrast between LABIAL and nonlabial (i.e. labiality) represented?

RQ3. How do place and labiality interact when both are contrastive in vowels?

For both tense and lax vowels we assess the same three phonological contrasts: single-feature contrasts with respect to either place (/ʏ/ - /ɔ/), single-feature contrast with respect to labiality (/ɛ/ - /ʏ/), and a two-feature contrast regarding both place and labiality (/ɛ/ - /ɔ/).

We determine whether or not a mispronounced prime still provides access to the canonical word form of the prime. For example, when looking at vowel pair /o-ø/, does a semantically related word with /o/ mispronounced as [ø] still prime the target and speed up reaction times compared to mispronunciations of unrelated words? So, if the word *boot* (/bot/ – 'boat') is mispronounced as [bøt]*, does it still prime target *schip* ('*ship*')? And vice versa, if a word with /ø/ is mispronounced as [o]? When within a single contrast one direction of change primes the target, but the opposite mispronunciation does not, this constitutes a perceptual asymmetry, which can inform us about the phonological representations in place.

We predict similar perceptual patterns as found in the EEG experiment in Chapter 2. We expect similar outcomes for both tense and lax vowels, as they constitute similar contrasts. We predict that correct pronunciations (CPs) of semantically related primes result in significant priming relative to unrelated CP primes, regardless of vowel condition. The predicted priming of mispronunciations (MPs) is given in Table 3.1 for tense vowels and Table 3.2 for lax vowels.

Regarding the place contrasts (tense /ø/ vs. /o/, and lax /ʏ/ vs. /ɔ/), we predict to find priming when a coronal is mispronounced as a dorsal, but not vice versa. The [DORSAL] SR feature will not mismatch with an underspecified coronal representation, and word access will take place, but if a dorsal vowel is mispronounced as a coronal vowel, mapping would result in a mismatch and hence no lexical access and no priming effect. Therefore, an asymmetry is predicted for this single-feature contrast.

Regarding labiality, we also predict an asymmetry. This prediction is relevant for tense /e/ vs. /ø/, and lax /ɛ/ vs. /ʏ/. We predict that when the feature [LABIAL] is extracted from the surface form of a semantically related MP, this will not result in lexical access of a word with a nonlabial vowel, because UR of a nonlabial vowel does not have a LABIAL node, which causes mapping of [LABIAL] SR feature onto a nonlabial UR to be aborted. This phonological conflict will hinder word access and hence will not result in semantic priming. However, a nonlabial MP will not trigger mapping, so this does not result in a phonological conflict. Word access is not hindered and semantic priming is predicted.

**Table 3.1** Predicted outcomes for Dutch tense mid vowels /e, ø, o/.

| Vowel pair | Contrastive dimensions | Featuraldistance | CP | MP | SR of MP | UR of CP | Mapping outcome | Priming | Asymmetry |
|---|---|---|---|---|---|---|---|---|---|
| **/ø/~/o/** | place | 1 | /ø/ | [o] | [DORSAL] | [Φ] | no-mismatch | Yes | YES |
| | | | /o/ | [ø] | [CORONAL] | [DORSAL] | mismatch | No | |
| **/e/~/ø/** | labiality | 1 | /e/ | [ø] | [LABIAL] | | *aborted* | No | YES |
| | | | /ø/ | [e] | | [LABIAL] | no-mismatch | Yes | |
| **/o/~/e/** | place & | 2 | /o/ | [e] | [CORONAL] | [DORSAL] | mismatch | No | NO |
| | labiality | | | | | [LABIAL] | | | |
| | | | /e/ | [o] | [DORSAL] | [Φ] | no-mismatch | No | |
| | | | | | [LABIAL] | | *aborted* | | |

CP = correct pronunciation. MP = mispronunciation. SR = surface representation. UR = Underlying representation. With priming, we mean positive priming, indicating word access by the mispronunciation. [Φ] indicates underspecification.

**Table 3.2** Predicted outcomes for Dutch lax mid vowels /ɛ, ʏ, ɔ/.

| Vowel pair | Contrastive dimensions | Featuraldistance | CP | MP | SR of MP | UR of CP | Mapping outcome | Priming | Asymmetry |
|---|---|---|---|---|---|---|---|---|---|
| **/ʏ/~/ ɔ /** | place | 1 | /ʏ/ | [ɔ] | [DORSAL] | [Φ] | no-mismatch | Yes | YES |
| | | | /ɔ/ | [ʏ] | [CORONAL] | [DORSAL] | mismatch | No | |
| **/ɛ/~/ʏ/** | labiality | 1 | /ɛ/ | [ʏ] | [LABIAL] | | *aborted* | No | YES |
| | | | /ʏ/ | [ɛ] | | [LABIAL] | no-mismatch | Yes | |
| **/ɔ/~/ɛ/** | place & labiality | 2 | /ɔ/ | [ɛ] | [CORONAL] | [DORSAL] [LABIAL] | mismatch | No | No |
| | | | /ɛ/ | [ɔ] | [DORSAL] [LABIAL] | [Φ] | no-mismatch *aborted* | No | |

CP = correct pronunciation. MP = mispronunciation. SR = surface representation. UR = Underlying representation. With priming, we mean positive priming, indicating word access by the mispronunciation. [Φ] indicates underspecification.

3

When both labiality and place features differ (i.e. vowel pairs /e-o/ and /ɛ-ɔ/), no priming is predicted in either direction of change. When a nonlabial coronal vowel is mispronounced as a labial dorsal vowel, there will be a conflict based on labiality: the labial SR feature triggers mapping, but this mapping is aborted since UR of a nonlabial vowel has no labial node for the labial feature to find a slot to be mapped onto. Vice versa, when a labial dorsal vowel is mispronounced as a nonlabial coronal vowel, the [CORONAL] SR feature will mismatch with the [DORSAL] feature in UR. As such, both MPs result in phonological conflict and no word access is predicted for either of them.

In sum, we predict asymmetries in single-feature contrasts. For the two-feature contrast no semantic priming is predicted for either of the MPs, as phonological conflict is predicted for both directions of change.

## Method

We ran a cross-modal semantic priming experiment. During the experiment, an auditory *prime* word is presented, followed by a written *target* word on screen. Subjects performed lexical decision on the target word through button presses. Primes are semantically related or unrelated to the target to induce semantic priming. Primes are correctly pronounced (CP) or mispronounced (MP) with regard to the vowel's labiality or place features, to assess the impact of such vowel mispronunciations on lexical access. In the current section, we will describe our subject pool, experiment design, stimulus materials, and experimental procedure in more detail.

### Subjects

46 adult native speakers of Dutch between 18 and 30 years old were included in the final sample (mean age = 21.9 years; SD = 2.6; 4 male). Subjects were recruited through Sona Systems – the Radboud Research Participation System, and were tested at the CLS (Centre for Language Studies) Lab at the Radboud University in Nijmegen (The Netherlands). Ethical approval for the study was obtained from the Ethische Toetsingscommissie Geesteswetenschappen of Radboud University (ETC-GW reference number 5523).

All subjects were born and raised in the Netherlands, and grew up in monolingual Dutch home-environments. All participants had normal hearing and normal or corrected-to-normal vision. Subjects had no speech or language impediments (current or in the past) or dyslexia, nor any other cognitive or neurological impairments.

Speakers of dialects with different vowel inventories than Standard Dutch or with morphophonological processes (such as *umlaut*) compared to Standard Dutch, such as

Achterhoeks and Limburgian dialects, were excluded from participation. People who grew up in Limburg were excluded from participating, because Limburg is a region where dialect is ubiquitous and even non-speakers will have been submerged to a large extent. Early exposure to a different vowel inventory and/or phonological processes may have impacted their vowel representations.

Two additional subjects were excluded from analysis due to speaking a Limburgian dialect (reported at test). We replaced them with two new subjects, aiming for 48 participants to be included. However, during analysis, two additional subjects were excluded for the following reasons: correct score < 75% (1 subject); reaction times > 2 SD slower than mean group RT (1 subject). Hence, data of 46 participants were included in the analysis.

Subjects signed an informed consent prior to their participation. They were reimbursed by means of gift vouchers or study credits.

**Design**

We used the semantic priming studies by Roberts et al., (2014); Kotzor et al. (2015), and Scharinger and Lahiri (2010) as a base to design our experiment.

The cross-modal semantic priming experiment had four main conditions regarding whether or not primes are semantically related to the target words and whether or not primes are correctly pronounced, and a set of fillers. In addition to these main conditions, there are twelve vowel conditions (six vowel pairs with two directions of change). Details on the vowel conditions follow after description of the four main conditions.

***Main conditions***

The cross-modal semantic priming experiment had the following conditions: (1) CP_related, (2) CP_control (i.e. unrelated), (3) MP_related, (4) MP_control (i.e. unrelated). Unrelated primes are used as a control condition to contrast with related primes. Conditions using unrelated primes are hence called control conditions.

In addition to the four conditions, we also used filler items. Conditions and fillers are listed with examples in Table 3.3. CP in condition (1) and (2) refers to *correct pronunciation* of the prime. MP in conditions (3) and (4) refers to *mispronunciation* of the prime. In MP conditions the stressed vowel of a CP word is substituted by an incorrect vowel, resulting in a nonword.

**Table 3.3** Overview of the four main conditions and the fillers, explained by means of an example of the /o/ [ø] vowel condition with CP /bot/ (= boat) and MP *[bøt].

| Condition | In words | Description of prime and target | Example stimuli | Trials per block |
|---|---|---|---|---|
| (1) CP_related | **Correctly pronounced** real word prime - semantically **related** to target | **Prime**:<br>• correctly pronounced real word<br>• semantically **related** to target | **Prime**:<br>*/bot/* (‘boat’) | 9 |
| | | **Target**: real word - synonym of prime | **Target**:<br>*schip* (‘ship’) | |
| (2) CP_control | **Correctly pronounced** real word prime - semantically **unrelated** to target | **Prime**:<br>• correctly pronounced real word<br>• semantically **unrelated** to target | **Prime**:<br>*/pozə/* (‘pose’) | 9 |
| | | **Target**: real word – unrelated to prime | **Target**:<br>*schip* (‘ship’) | |
| (3) MP_related | **Mispronounced** prime - semantically **related** to target | **Prime**:<br>• mispronounced word: vowel is substituted<br>• meaning of correctly pronounced version is semantically **related** to target. | **Prime**:<br>*[bøt]** | 9 |
| | | **Target**: real word - synonym of prime | **Target**:<br>*schip* (‘ship’) | |
| (4) MP_control | **Mispronounced** prime - semantically **unrelated** target | **Prime**:<br>• mispronounced word: vowel is substituted<br>• meaning of correctly pronounced version is semantically **unrelated** to target. | **Prime**:<br>*[pøzə]** | 9 |
| | | **Target**: real word – unrelated to prime | **Target**:<br>*schip* (‘ship’) | |
| (5) Filler | Nonword targets - filler items to ensure proper lexical decision task – not an experimental condition | **Prime**: Correctly pronounced words (CPs) or their mispronunciations (MPs) | **Prime**:<br>*/bot/* (‘boat’)<br>or *[bøt]** | 18 |
| | | **Target**: nonword | **Target**: *mindaan** | 18 |
| **Total** | | | | 72 |

Prime is an auditory word presented through headphones. Target is a written word presented on screen.

In the *related* conditions (1) CP_related and (3) MP_related, paired prime and target have a (near-)synonym semantic relation; for example, prime *boot* (= boat) and target *schip* (= ship). The CP_related condition is expected to elicit a priming effect compared to the CP_control condition where prime and target are not related. In the MP_related condition, a pair of synonyms (prime and target) is presented, but the prime is mispronounced. Hence, the prime is a nonword without actual semantic content.



In the *unrelated* conditions (2) CP_control and (4) MP_control, prime and target have no semantic nor any other relationship. These unrelated conditions serve as control conditions (i.e. baselines) in order to calculate priming effects for conditions (1) CP_related and (3) MP_related respectively. Priming effects are calculated by comparing related and unrelated primes.

Lastly, filler trials were included, which were trials with nonword targets. These fillers ensured actual lexical decision. Participants saw 50% real word targets and 50% nonword targets. This was used to calculate % correct scores to assess task performance per subject. Fillers are not relevant to our research questions and are not included in the final analysis.

***Vowel conditions***

In the current experiment we assess processing of three tense vowels /e, o, ø/, and three lax vowels /ɛ, ɔ, ʏ/. Each vowel is paired with both other vowels within its set, forming a phonological contrast of interest, resulting in six vowel pairs. Each vowel pair has two directions of change, resulting in 12 vowel conditions (see Table 3.4). These two orders are required to investigate perceptual asymmetries: we can measure MP priming in both directions of a contrast.

For example, vowel condition (5) /o/ - [e] with target word *huis* (= house) has CP prime *w[o]ning* (= house) with MP **w[e]ning*, and vice versa: vowel condition (6) /e/ - [o] with target word *angst* (= fear) has CP vr[e]s (= *fear*) with MP *vr[o]s.

**Table 3.4** Overview of 6 vowel pairs and 12 vowel conditions.

| Vowel pair | Contrast | Vowel condition | Features of interest of prime | |
|---|---|---|---|---|
| | | | CP | MP |
| **(1) e - o** | place | (1) /ø/ - [o] | CORONAL | DORSAL |
| | | (2) /o/- [ø] | DORSAL | CORONAL |
| **(2) e - ø** | labiality | (3) /e/- [ø] | nonlabial | LABIAL |
| | | (4) /ø/ - [e] | LABIAL | nonlabial |
| **(3) o - ø** | place & labiality | (5) /o/ - [e] | DORSAL LABIAL | CORONAL |
| | | (6) /e/ - [o] | CORONAL | DORSAL LABIAL |
| **(4) ɛ - ɔ** | place | (7) /ʏ/ - [ɔ] | CORONAL | DORSAL |
| | | (8) /ɔ/ - [ʏ] | DORSAL | CORONAL |
| **(5) ɛ - ʏ** | labiality | (9) /ɛ/ - [ʏ] | nonlabial | LABIAL |
| | | (10) /ʏ/ - [ɛ] | LABIAL | nonlabial |
| **(6) ɔ - ʏ** | place & labiality | (11) /ɔ/ - [ɛ] | DORSAL LABIAL | CORONAL |
| | | (12) /ɛ/- [ɔ] | CORONAL | DORSAL LABIAL |

Tense vowels in upper part (vowel pairs 1, 2, and 3). Lax vowels in lower part (vowel pairs 4, 5, and 6). /V/ indicates vowel in CP, [V] indicates vowel in MP. Vowel condition hence indicates vowel change. Phonological features are indicated in SMALL CAPITALS.

### Stimuli

36 items (i.e. pair of prime + target) were created for each main condition (i.e. CP_related, CP_unrelated, MP_related, MP_unrelated) for all 12 vowel conditions. For each vowel condition we used:

(1) 36 semantically related prime-target pairs (both a CP and MP variant) for conditions (1) CP_related and (3) MP_related;

(1) 36 semantically unrelated (control) prime-target pairs for conditions (2) CP_control and (4) MP_control. Target as well as prime words are identical to (1), only paired differently;

In addition, there were 36 nonword targets paired with CP and MP primes as fillers.
Primes were words presented auditorily. Primes were recorded by a male native Dutch speaker using recording software Audacity (version 1.2.4) in a sound attenuated studio booth with a microphone of professional quality. For each prime a single token was used in the experiment.

Intensity was normalized to 70 dB in acoustic analysis programme PRAAT (Boersma & Weenink, 2012).

Ideally, primes should be only nouns and only words with similar length (preferably monosyllabic) and morphology (preferably monomorphemic). Furthermore, they should have synonyms (as targets) and they should result in nonwords when mispronounced. These criteria resulted in too few suitable stimuli. We therefore decided to be more tolerant regarding parts of speech, word length and morphology, provided that all variation or deviations from the most conservative terms were distributed equally over all vowel conditions.

Only content words were included as stimuli (nouns, verbs, adjectives and adverbs), because their semantic content is required in semantic priming. Primes had synonyms to be used as targets. Frequency of primes was matched across vowel conditions based on frequencies in SUBTLEX-NL database (Keuleers, Brysbaert & New, 2010). Primes were one to three syllables long. Prime length in syllables was matched across conditions. The critical vowels /e, o, ø/ and /ɛ, ɔ, ʏ/ appeared in stressed position. All primes started with a consonant. Words were mono- or dimorphemic. To create the MP conditions, all CP primes had to become nonwords when the critical vowel was substituted.

Semantically related primes and targets were (near-)synonyms. Targets were one to four syllables long. Prime and target were not phonologically related. Paired primes and targets always were the same parts of speech.

Similar to Roberts et al. (2014) we ran a familiarity judgement task (n = 63) as well as a semantic similarity judgement task (n = 75) to assess whether all words were familiar, and whether synonym relations were sufficiently strong for semantically related prime-target pairs, and sufficiently weak for controls (see Appendix 5).

Apart from synonym pairs, the experimental design also requires unrelated prime and target pairs as a control condition. Since it was infeasible to find sufficient prime-target pairs fitting all requirements, we used each prime both as a related prime (paired with synonym target) and an unrelated prime (paired with a semantically unrelated target).

Filler trials had nonword targets that varied in length similarly to the other targets across conditions.

**Procedure**
We conducted a cross-modal semantic priming experiment using Presentation Software (version 18.2 02.18.16, www.neurobs.com). The experiment was conducted in a sound attenuated air-conditioned booth where participants were seated in a comfortable chair with a PC display on a desk approximately 1 m in front of them. Chair height was adjusted to screen centre at eye level.

In every trial, an auditory *prime* word was presented, followed by a written *target* word on screen, for example [bot] (*boat*) followed by written target word 'schip' (*ship*). Primes were presented via over-ear Sennheiser headphones. Primes were presented at a comfortable intensity of 70 dB SPL. Loudness was identical for all subjects.

The target words appeared as a written word on screen centre, in lower case letters, after an ISI of 250 ms after the offset of the auditory prime, and was presented for 300 ms, similar to the paradigm used in Roberts et al. (2014). Subjects performed a lexical decision task based on the target word by means of button presses on a button box. The maximum response time (timeout) was 2000 ms, similar to Scharinger & Lahiri (2010). The next trial started 800 ms after a button press or timeout. Figure 3.1 depicts the timeline for a single trial.

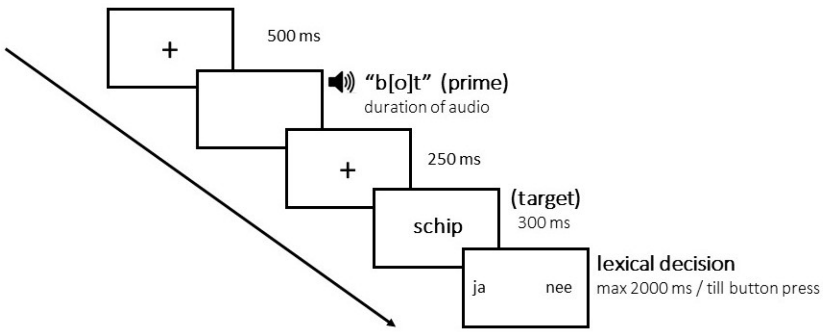



**Figure 3.1** Overview of a single trial.

Subjects were instructed to respond to the target by judging whether or not that word exists in Dutch by pressing a button corresponding to their answer on a button box: ja (*yes*) or nee (*no*). We used balanced response mapping: 50% of participants pressed YES (= real word) with their dominant index finger and NO with the index finger of the other hand. The other half pressed YES with their non-dominant hand. Similar to Roberts et al. (2014), this balanced response mapping was used to compensate for response preparation effects in the contralateral hemisphere

(Kutas & Donchin, 1977). Participants were instructed to respond as quickly and accurately as possible. After oral instruction, participants could read the instructions once again on screen before starting the experiment.

The experiment consisted of one training block (12 trials) to familiarize the participants with the procedure, followed by 12 experimental blocks (72 trials each). During training, trials were similar, but non-identical to the actual test trials. All vowel pairs, correct and incorrect primes, and real word as well as nonword targets appeared in the practice block. The only difference compared to the experimental blocks was that subjects received feedback on their performance. At the end of each training trial, a green happy smiley-face indicated a correct and sufficiently fast response. A red sad smiley-face indicated an incorrect response or a response that was not swift enough (> 2000 ms after target offset). After training, participants could ask the experimenter clarification questions before starting the experimental blocks.



A Latin Square design was used, resulting in twelve blocks of 72 trials per participant. Each target occurred once per participant. All vowel conditions occurred in all blocks. In between blocks subjects took breaks to reduce fatigue. The experiment proceeded as soon as the participant indicated that he/she was ready for the next block by means of a button press. The complete experiment – including instruction, training, and post-test questionnaire - lasted roughly 45 minutes.

We used four stimulus lists with randomized block orders and randomized order of stimuli within blocks. Which items were presented in which block was established in pseudorandomized lists of prime-target pairs for each of the four versions. Each prime-target combination was assigned to one out of the four different list versions of the experiment. Participants were randomly assigned to one of four list versions. In each block a participant saw 36 real-word targets and 36 nonword targets.

After the cross-modal priming experiment, the participants filled out a digital questionnaire on a laptop using Castor EDC (2015). This questionnaire consisted of three parts:

(1) Personal information: questions concerning age, handedness, and language background;
(2) Familiarity judgement task;
(3) Semantic similarity judgement task.

Part 1 was used to verify whether participants met all inclusion criteria, and whether the correct experiment version with respect to response sides and handedness was conducted. Part 2 and 3 were similar to the judgement tasks used prior to running the experiment (see Appendix 5).

Items that were changed based on the initial judgement tasks, were included in the post-test questionnaire for post-hoc verification. Stimuli not meeting the requirements were not included in final analysis (see Appendix 5).

## Data cleaning

Participants with < 75% correct scores were excluded from further analysis (1 subject). Items that had low familiarity scores (2 items) or inadequate semantic similarity scores (8 items) in the post-experiment familiarity judgement questionnaire were excluded from further analysis.

Trials with RTs < 300 ms were excluded from analysis. Data points more than 3 SD away from a subject's mean were defined as outliers. These were excluded from further analysis. Reaction times above 1500 ms were excluded from further analysis, as after that time reaction times will not reflect initial lexical access only[4]. Nearly all RTs over 1500 ms belonged to a single subject. Most of this subject's responses exceeded 2 SD of group mean RT. This subject was excluded from further analysis. Overall, CP priming was confirmed before any further steps in data-cleaning.

Only items that yield semantic priming are informative when investigating how mispronunciations impact lexical access. If an MP does not prime, but its CP counterpart does not prime either, the lack of MP priming cannot be interpreted in a meaningful way. Hence, non-priming prime-target pairs were excluded from analysis. As a first step in data cleaning, we calculated the degree of priming for every prime-target pair by comparing reaction times for a related CP prime and an unrelated (= control) CP prime. Only items with a shorter RT for related than unrelated CP primes remained included. Based on this requirement, we eliminated roughly 30% of our data.

Why a target is not primed is not always clear. Non-priming could occur regardless of adequate familiarity and semantic similarity scores during pretests. Non-priming could also be caused for example by control and related primes that are not well-matched (e.g. with respect to word length and/or morphology). If the morphology of the related prime is more complex, this may slow reaction times down. Since the control and related primes did not always match in all respects, this may have impacted the presence or absence of priming per individual target.

### Extra cleaning considerations

Apart from the planned cleaning steps mentioned already, we had to reconsider some of our decisions regarding stimulus criteria. The number of items included in the final analysis is

4 Personal correspondence with Prof. dr. Henning Reetz (University of Oxford), November 2018.

reported in the Analysis section of the current chapter. We do not split the excluded items out by each criterion, as many of the excluded items did not meet multiple requirements.

***Word class***
Strong semantic content is required to yield priming effects. Depictable nouns are preferred in semantic priming studies for primes as well as targets, as they typically provide the strongest priming. Due to a lack of available items, we also included other content words, which we expected to have enough semantic content to yield semantic priming. About half of our items were nouns.



In retrospect, including word classes other than nouns may not have been ideal, as data showed that priming was particularly found for nouns. One reason for nouns tending to prime stronger is because verbs and adjectives/adverbs largely depend on their context to be interpreted. Their meaning is often relative (e.g. words like *large*). Nouns tend to have a stronger and more straightforward semantic content, less dependent on sentence context.

In order to determine whether it would indeed be sensible to include only nouns, we examined the impact of word class on degree of priming in our dataset (regardless of vowel condition). We found a significant priming effect for all word classes, but there was more variance in verbs and adjective/adverbs than in nouns, and the degree of priming was three times stronger in nouns. Consequently, in order to have a measure sensitive enough to potentially reveal asymmetries in priming we decided to include the strongest and most consistent word class, i.e. nouns. Other word classes would introduce more variance, and reduce priming effects.

***Considerations for primes***
Ideally, critical vowels occur in a stressed position, because stressed vowels are never reduced and are more important in word recognition. Ideally, the stressed vowel is in the first syllable of the word. At the position of the first syllable, mispronunciations often result in more probable word forms than later in words – their uniqueness point may be later in the word than the critical vowel. For a mispronunciation to be taken as a potential different word form (or pseudoword), it should result in a probable word form. Some vowels rarely occur in later syllables and/or in particular consonantal contexts. Also, morphology impacts probability for a vowel to occur, e.g. as part of a suffix.

We initially included words with non-initial stress or with the critical vowel in a later position. Post hoc inspection showed that the uniqueness point of these words is often before the critical vowel. In such cases, lexical access occurs before the end of the word is reached and hence regardless of whether or not the vowel is mispronounced. Consequently, analysis of

MP priming becomes meaningless. For this reason, we excluded items without initial stress or critical vowel at the initial syllable from further analysis.

Ideally, stimuli should all be monomorphemic and monosyllabic. However, due to the limited number of items available, we also included longer or more complex items. Post hoc, we decided to exclude primes longer than two syllables (and their paired targets) for a number of reasons. First, longer words provide more information irrespective of the mispronunciation that may enable lexical access. This may for example be the case for the prime *t[o]venaar* ('wizard'). Mispronunciations as **t[e]venaar* or **t[ø]venaar* still have two syllables *venaar* following the mispronounced vowel. These final syllables occur in only two Dutch words: *tovenaar* or *evenaar* ('equator'). As there are hardly any competitors, the word is likely to activate *tovenaar* irrespective of a mispronunciation in the initial vowel. Also, compound(like) nouns like *wedstrijd* (game/contest), have a strong second syllable which may retroactively disambiguate the word, because the second syllable does generally not occur without the first and hence will provide lexical access regardless of MP. Such words were excluded from the data.

Moreover, related and control primes that belong to the same target should be matched regarding length and morphological makeup. We did balance the occurrence of different kinds of words (e.g. regarding length or morphological makeup) over vowel categories and we did match word classes, but related and unrelated primes that belonged to a single target were not always similar in number of syllables or morphemes. Both the control and a related prime for a particular target are required in order to calculate the degree of priming. Post hoc inclusion of only monosyllabic/monomorphemic words would thus result in excluding (nearly) all targets. To leave sufficient data to analyse, both mono- and disyllabic word were included, but longer words were excluded.

Lastly, we reconsidered primes with tense vowels followed by liquids. If /ø/, /o/ or /e/ are followed by an /r/ or /l/ (closed syllable), as for example in *kleur*, *deur*, or *geul*, the vowel is realized as /ʏː/, /ɔː/, or /ɪː/ respectively, because of effects of phonological context on phonetic realization of the vowel categories. We initially included these, assuming no differences at a phonological level with occurrences in a different context. However, these vowels are qualitatively different in surface form and might even be lax vowels in the representation. Thus, in order to compare tense and lax vowels, tense vowels followed by liquids were excluded from further analysis.

***Considerations for targets***

A target was excluded if either one of its paired primes was excluded due to any of the above-mentioned criteria, because both related and control primes are required in order to determine degree of priming for a particular item. Similarly, targets that lacked data in one or both directions - for instance when no participant had a correct and timely response – were also excluded (8 targets).

Target words which are long or morphologically complex may need more processing time, or require more than one fixation. We excluded morphologically complex words. As we did not find an effect of target word length (in letters) on reaction time, we did not set any extra post hoc restrictions on target length.



## Analysis

58 targets were included in the final analysis. The number of targets per condition is reported in Table 3.5 for tense vowels and in Table 3.6 for lax vowels. Only real word target trials are included in the final analysis.

**Table 3.5** Number of targets and trials per condition for tense vowels.

| Vowel pair | Vowel condition | Contrastive features | CP | MP | No. of trials | No. of targets |
|---|---|---|---|---|---|---|
| /o- ø/ | /o/ [ø] | place | /o/ | [ø] | 417 | 10 |
| | /ø/ [o] | | /ø/ | [o] | 128 | 3 |
| /e- ø/ | /e/ [ø] | labiality | /e/ | [ø] | 211 | 5 |
| | /ø/ [e] | | /ø/ | [e] | 363 | 9 |
| /o-e/ | /e/ [o] | place & labiality | /o/ | [e] | 151 | 4 |
| | /o/ [e] | | /e/ | [o] | 207 | 5 |

**Table 3.6** Number of targets and trials per condition for lax vowels.

| Vowel pair | Vowel condition | Contrastive features | CP | MP | No. of trials | No. of targets |
|---|---|---|---|---|---|---|
| /ɔ - ʏ/ | /ɔ/ [ʏ] | place | /ɔ/ | [ʏ] | 89 | 2 |
| | /ʏ/ [ɔ] | | /ʏ/ | [ɔ] | 74 | 2 |
| /ɛ - ʏ/ | /ɛ/ [ʏ] | labiality | /ɛ/ | [ʏ] | 222 | 5 |
| | /ʏ/ [ɛ] | | /ʏ/ | [ɛ] | 78 | 2 |
| /ɔ -ɛ/ | /ɛ/ [ɔ] | place & labiality | /ɛ/ | [ɔ] | 205 | 5 |
| | /ɔ/ [ɛ] | | /ɔ/ | [ɛ] | 230 | 6 |

Reaction times were analysed with statistics software JMP (SAS, 2018) using linear mixed effect model with fixed factors *Relatedness* (2 levels: related, control), *Wordness* (2 levels: CP or MP), and *Vowel Condition* (12 levels), with random factors *Subject* and *Item* (i.e. target words). In an initial model, we also included fixed factor *Tenseness* (2 levels: tense/lax)[5]. Since no main effect of *Tenseness* was found, it was removed from the model. Tense as well as lax vowels were included in the final model in order to have an optimal estimate of random factor *Subject*. Both random intercepts and random slopes were used. Furthermore, we used planned comparisons within this model.

To determine whether an MP hinders lexical access, we tested whether MP priming differs from CP priming in that particular vowel condition. If MP priming is significantly less than CP priming, MP still hinders lexical access. To determine whether the two directions within one vowel pair are asymmetrical, we intended to test whether opposite vowel conditions (i.e. two directions within one vowel pair) differ with respect to the difference in priming for MPs and CPs. However, since after data cleaning only a limited dataset remained, our data lacked power to test asymmetries this way.

As we predicted asymmetries in all single-feature contrasts and also considered potential asymmetries in two-feature contrasts for both tense and lax vowels, we were interested to see whether or not mispronunciations yielded priming in all vowel conditions. Hence, we tested whether there was CP priming and whether there was MP priming for all twelve vowel conditions as planned comparisons. We contrasted CP control and CP related trials: significant difference between the two would mean significant priming, with positive priming indicating related primes to yield shorter reaction times than unrelated primes. Similarly, for MPs we calculated contrasts between MP control and MP related trials. Outcomes are given in the Results section below.

Due to our limited dataset, we evaluate patterns rather than significant asymmetries, as trends can still be informative.

## Results

In the initial model where we included *Tenseness*, we found no main effect for *tenseness* (tense vs. lax, $F(1,54.61) = 0.243$, $p = 0.63$) nor any interactions with *Relatedness* or *Wordness*, after which we removed this factor from the model.

There was a main effect of *Relatedness* (control vs. related primes), $F(1,2253) = 36.98$, $p < 0.0001$, indicating that related primes yield priming. Reaction times for related trials were 34 ms shorter than for control trials. Note that the main effect of *Relatedness* was also significant without data cleaning.

5 The factor Tenseness is biased with Vowel Condition, so for this test Vowel Condition was not included as a factor.

Reaction times to MPs were 11 ms slower than to CPs. This difference was not significant: there was no main effect of *Wordness* (CP vs. MP), $F(1,2248) = 2.6$, $p = 0.107$.

There was a significant interaction of *Relatedness*Wordness*, $F(1,2256) = 11.53$, $p = 0.0007$, indicating different degrees of priming for CPs and MPs. CPs had 52 ms priming (RT control (LSM) = 594 ms; RT related (LSM) = 541 ms), whereas MPs had 15 ms of overall priming (RT (LSM) control = 584 ms; RT (LSM) related = 568 ms).

3

No main effect of *Vowel Condition*, $F(11,45.02) = 0.62$, $p = 0.8$, nor any further interactions with *Vowel Condition* were revealed.

Planned comparisons are reported in Table 3.7 for tense vowels and Table 3.8 for lax vowels. The limited amount of data reduces sensitivity for all conditions in the model, since less data results in less accurate estimates of random factors. We may lack data for CP priming per vowel to reach significance even in case of large RT differences between related and control primes. Results should be interpreted with caution.

Below, we first describe our findings for single-feature contrasts (contrasts of place and labiality), followed by our findings for the two-feature contrasts.

**Place contrasts**

Results for the tense vowel pair /o-ø/ show significant priming for both CPs: 38 ms of priming for /o/, and 74 ms of priming for /ø/. There was less priming when the dorsal was mispronounced as coronal (-14 ms) than vice versa (35 ms) (n.s.). Thus, /o-ø/ vowel pair might show an asymmetrical pattern in MP priming.

Lax vowels in the vowel pair /ɔ-ʏ/ show positive priming for both CPs, but only CPs of /ʏ/ show significant priming (53 ms). Similar to the tense vowels, there was numerically less priming when the dorsal was mispronounced as coronal (5 ms) than vice versa (20 ms). However, since CPs of /ʏ/ did not prime, it is not clear what priming of words with an MP of /ʏ/ actually means.

**Table 3.7** Results of priming analyses per condition for tense vowels – contrasting RT control (i.e. unrelated) and RT related.

| **Vowel pair** | **Vowel condition** | **Contrastive features** | **CP** | | | | | **MP** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | **priming (SD)** | **RT control (SD)** | **RT related (SD)** | **t-ratio** | **p-value** | **priming (SD)** | **RT control (SD)** | **RT related (SD)** | **t-ratio** | **p-value** |
| /o- ø/ | /o/ [ø] | place | 38.3 (15.8) | 602.5 (21.9) | 564.2 (22.0) | 2.42 | **0.015** | -14.4 (16.3) | 568.6 (29.3) | 583.0 (22.1) | -0.88 | 0.37 |
| | /ø/ [o] | | 74.2 (29.3) | 606.1 (38.4) | 531.9 (36.6) | 2.52 | **0.011** | 35.2 (28.9) | 596.4 (37.4) | 561.2 (37.2) | 1.21 | 0.22 |
| /e- ø/ | /e/ [ø] | labiality | 25.2 (23,4) | 559.5 (30.3) | 534.4 (29.4) | 1.07 | 0.28 | 26.8 (22.7) | 575.8 (29.6) | 548.9 (29.8) | 1.18 | 0.24 |
| | /ø/ [e] | | 54.6 (17.4) | 599.7 (23.2) | 545.1 (23.1) | 3.13 | **0.0017** | 16.4 (17.6) | 569.7 (23.7) | 553.3 (22.9) | 0.92 | 0.35 |
| /o-e/ | /e/ [o] | place & labiality | 56.3 (26.1) | 599.6 (33.2) | 543.3 (33.3) | 2.15 | **0.03** | 39.4 (26.6) | 635.9 (33.4) | 596.5 (33.5) | 1.48 | 0.14 |
| | /o/ [e] | | 61.7 (23.9) | 596.6 (30.6) | 534.9 (29.5) | 2.58 | **0.0099** | -7.2 | 575.3 (29.3) | 582.5 (29.6) | -0.33 | 0.74 |

Priming is reported in milliseconds. Reaction times (RT) are least square means. Significant p-values are **bold.**

**Table 3.8** Results of priming analyses per condition for lax vowels – contrasting RT control (i.e. unrelated) and RT related.

| Vowel pair | Vowel condition | Contrastive features | CP | | | | | MP | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | priming (SD) | RT control (SD) | RT related (SD) | t-ratio | p-value | priming (SD) | RT control (SD) | RT related (SD) | t-ratio | p-value |
| /ɔ - ʏ/ | /ɔ/ [ʏ] | place | 33.8 (23.3) | 588.0 (30.4) | 554.2 (29.2) | 1.45 | 0.147 | 5.84 (23.3) | 582.9 (30.1) | 577.1 (29.5) | 0.25 | 0.80 |
| | /ʏ/ [ɔ] | | 53.7 (21.9) | 595.9 (28.3) | 542.3 (27.6) | 2.45 | **0.014** | 20.5 (21.6) | 591.2 (28.2) | 570.7 (27.2) | 0.95 | 0.34 |
| /ɛ - ʏ/ | /ɛ/ [ʏ] | labiality | 42.4 (22.4) | 536.9 (29.3) | 494.5 (29.6) | 1.89 | **0.058** | -19.2 (21.9) | 533.3 (29.1) | 552.5 (29.6) | -0.87 | 0.38 |
| | /ʏ/ [ɛ] | | 51.8 (36.5) | 618.2 (46.3) | 566.3 (44.9) | 1.42 | 0.15 | 47.6 (37.7) | 590.1 (46.4) | 542.5 (46.0) | 1.26 | 0.207 |
| /ɔ -ɛ/ | /ɛ/ [ɔ] | Place & labiality | 75.5 (34.7) | 581.1 (44.5) | 505.5 (45.3) | 2.17 | **0.0298** | 22.3 (35.1) | 549.7 (45.3) | 527.3 (44.8) | 0.63 | 0.52 |
| | /ɔ/ [ɛ] | | 65.3 (39.5) | 641.4 (47.5) | 576.1 (46.1) | 1.65 | **0.09** | 3.9 (36.9) | 635.7 (45.3) | 631.7 (46.5) | 0.11 | 0.91 |

Priming is reported in milliseconds. Reaction times (RT) are least square means. Significant p-values are **bold.**

3

**Labiality contrasts**

For the tense vowel pair /ø-e/, we find significant CP priming for words with /ø/ (55 ms), but not for CPs with /e/ (25 ms). Hence priming values for when nonlabial /e/ is mispronounced as labial, are not meaningful. When labial vowel /ø/ is mispronounced as nonlabial, priming is less (16 ms) than in CP. The data does not provide enough information to state anything about any asymmetry for this vowel pair.

For the labiality contrast in the lax vowels /ɛ-ʏ/ there is significant priming for both CPs (42 and 58 ms respectively). The data suggests an asymmetrical pattern of MP priming for the vowel pair /ɛ-ʏ/, with less semantic priming when nonlabial vowel is mispronounced as labial (-19 ms) than vice versa (48 ms).

**Two-feature contrasts**

Results for the tense vowel pair /e,o/ show significant priming for both CPs; 56 ms of priming for CPs with /e/ and 61 ms of priming for CPs with /o/. Vowel pair /e-o/ shows an asymmetrical pattern in MP priming: a mispronunciation of the dorsal (labial) vowel as coronal yields less semantic priming (-7 ms) than vice versa (39 ms).

Lax vowels /ɛ-ɔ/ show significant priming for CPs with /e/ (76 ms), but no significance is reached for CPs with /ɔ/, despite the large amount of priming of 65 ms ($p = 0.09$). We expect this to be due to large variance and small sample sizes. There is more priming when coronal vowel /ɛ/ is mispronounced as [ɔ] (22 ms) than vice versa (4 ms). However, due to a lack of CP priming, this is rather inconclusive.

## Discussion

We conducted a semantic priming experiment as a follow up to the EEG study in the previous chapter, investigating how place and labiality features are represented in Dutch vowels at a lexical level. The study had two aims. First, we were interested to find out whether the perceptual asymmetries reported in Chapter 2 would be replicated in the current study. Second, by including not only tense but also lax vowels, we aimed to find out whether diphthongization plays a role in the way vowels are represented in Dutch listeners.

Unfortunately, essential data cleaning resulted in a limited dataset. Therefore, it was not feasible to run reliable statistical tests to assess asymmetries. However, we did evaluate priming patterns in correct and mispronounced words.

Based on our hypothesized feature geometry and mapping process as proposed in Chapter 2, we predicted similar perceptual asymmetries in tense and lax vowels for single-feature contrasts (place or labiality). Regarding place contrasts, we predicted semantic priming – indicating lexical access – when a coronal vowel is mispronounced as a dorsal vowel, but not vice versa. Regarding labiality contrasts, we predicted semantic priming when a labial vowel is mispronounced as a nonlabial vowel, but not vice versa. For two-feature contrasts (labiality and place are contrastive) we predicted no semantic priming for mispronunciations in either direction, because both directions presumably constitute phonological conflicts, either based on place or on labiality. In sum, we predicted similar outcomes as reported in Chapter 2.



**Current findings**

Generally, we predicted correctly pronounced primes to prime semantically related target words. Without this CP priming, studying MP priming is meaningless. Indeed, we found overall priming for correctly pronounced words, indicating the method worked to induce semantic priming effects. We found significant positive CP priming in ten out of twelve vowel conditions. Two vowel conditions showed near-significant results (conditions /e/[ø] and /ɔ/[ʏ]). Despite non-significance, these vowel conditions all did yield a clear degree of positive priming, which reassures us in our assumption that these conditions would reach significance with a larger amount of data.

In general, none of the mispronunciations yielded significant priming effects. Since none of the MPs primed significantly, we cannot conclude any of the MPs to provide lexical access, nor could any (significant) asymmetries be revealed. However, we will describe the priming patterns we examined for the three contrasts of interest in both tense and lax vowels.

Predicted asymmetries for the single-feature contrasts and observed patterns in current experiment for tense vowels /e, ø, o/ and lax vowels /ɛ, ɔ, ʏ/ are presented in Table 3.9. Overall, the predicted coronal-dorsal asymmetry was observed both in tense and lax vowels. Regarding the labiality contrast, however, results are less clear. The predicted labiality asymmetry seemed present for lax vowels, but no clear pattern could be determined for tense vowels. For the two-feature contrast, no asymmetry was predicted, because either direction of change was predicted to result in phonological conflict. Tense vowels seemed to show an asymmetrical pattern in the same direction as the coronal-dorsal asymmetry attested for the place contrast. Results for lax vowels were unclear.

**Table 3.9** Overview of predictions based on Chapter 2 and current outcomes for tense and lax vowels in semantic priming.

| | | place | labiality | two-feature |
|---|---|---|---|---|
| Outcomes | Tense vowels | dorsal → coronal | ? | dorsal → coronal |
| | Lax vowels | dorsal → coronal | nonlabial → labial | ? |
| Predicted asymmetries | | dorsal → coronal | nonlabial → labial | no asymmetry |

Arrows indicate the direction of change perceptually most prominent, i.e. less priming than vice versa. Outcomes reflect observed patterns rather than significant asymmetries. The grey cells indicate outcomes different from what was predicted.

In sum, current priming patterns appear in line with underspecification of coronal and specification of dorsal in Dutch words in both tense and lax vowels. For the place contrast, findings from Chapter 2 seem to generalize to this different method. Data remain inconclusive regarding representations of labiality.

The statistical model did not yield a main effect of, or interactions with, tenseness, thus did not provide evidence for a difference between tense and lax vowels. Hence, we cannot state that diphthongization plays a role in representing these vowels at a phonological level.

When glancing over the results for tense and lax vowels, the labiality contrasts appear potentially different for tense versus lax vowels. Although this may be related to a difference in degree of diphthongization, a potential difference in height between /ʏ/ and /ø/ could also be considered. As discussed in Chapter 1, there is no consensus regarding height in particular with respect to /ʏ/ (compare Moulton, 1962; Zwaardemaker & Eijkman, 1928; Booij, 1989; Rietveld et al., 2004). It is hard to resolve this lack of agreement, since /ʏ/ can be distinguished from all other vowels either as a high, or as a mid-vowel. For vowels /ɛ/ and /ɔ/ opposite asymmetrical predictions would be made regarding discrimination of lax vowels depending on whether MID is regarded an underspecified or not specified feature (see Chapter 5 for more details on this matter). We conclude that current data cannot resolve the issue of the placement of /ʏ/.

**Limitations**

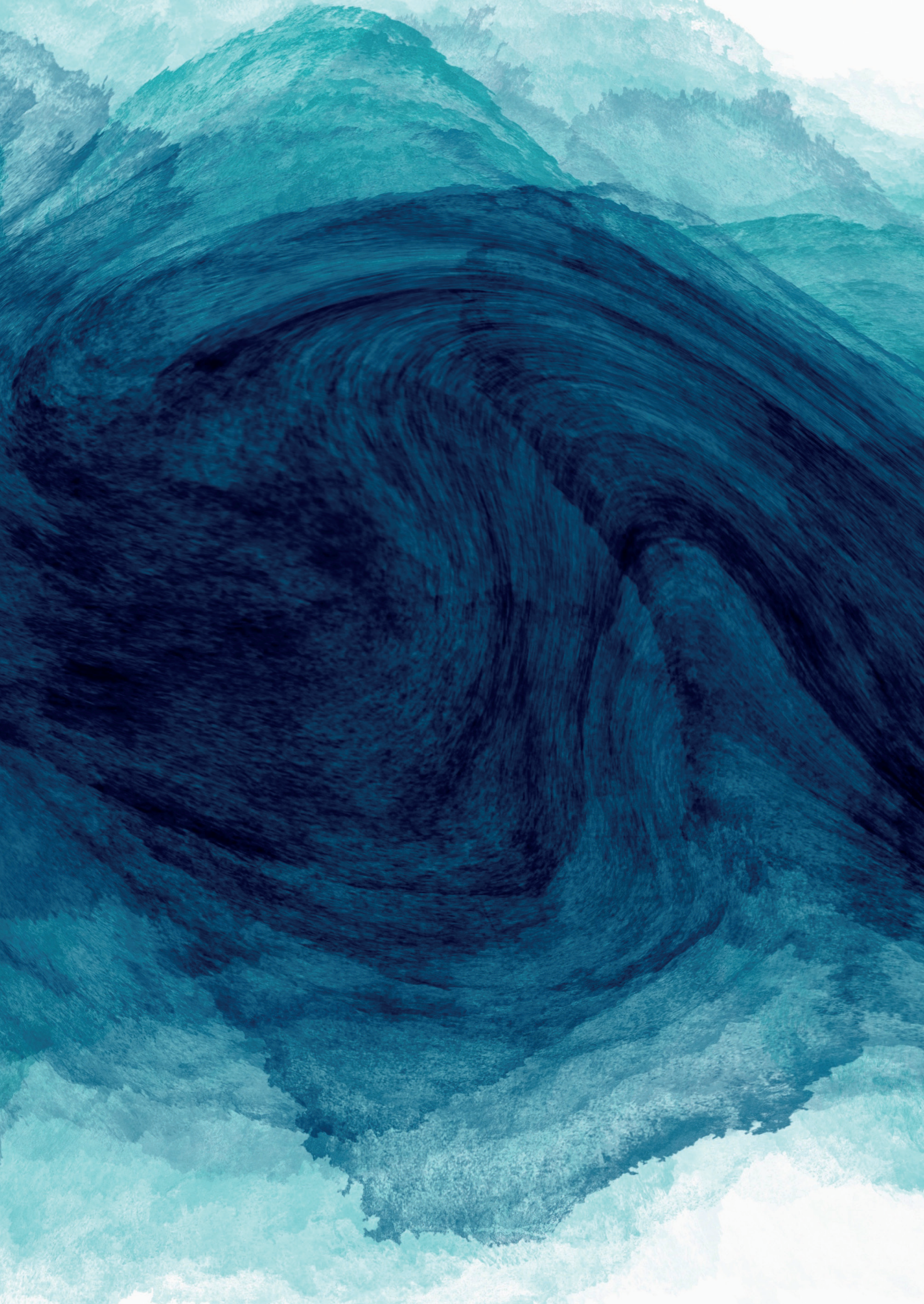
A number of our test items did not fit our adapted criteria to result in meaningful data (see the Data Cleaning section). We did try a post hoc attempt to create a new item list with stimuli that met all requirements. However, this resulted in an extremely limited list of items, in particular for conditions with front rounded vowels, more so for lax vowels, since those vowels are less frequent and also rarely occur in words that result in nonwords when the vowel is substituted by a nonlabial or back counterpart. Hence, it became a matter of testing a limited number of items with lots of participants, or a matter of loosening the criteria. In the current experiment we choose the latter, but in retrospect we realized that it would have been better to run many more participants on a short but clean list of items. Unfortunately, it was not feasible to repeat this within the scope of the current project.

Alternatively, an ERP experiment measuring N400 (similar to to Friedrich, Eulitz, & Lahiri, 2006), or a speeded auditory lexical decision task (similar to Lahiri and Reetz, 2010) could be considered. Although partly similar stimulus restrictions apply when using either of these paradigms, neither requires semantically related word pairs, enlarging the pool of potential test items.



**Conclusion**

Our findings support underspecification rather than full specification of phonological features, since we found asymmetrical patterns in both tense and lax vowels. Regarding place, results are in line with Chapter 2, but findings are unclear regarding labiality. The paradigm used in the current experiment appeared unsuitable to investigate place and labiality vowel contrasts in Dutch. Based on current data we cannot confirm nor reject the hypothesized vowel representations. A follow-up study would be required, such as an N400 or speeded auditory lexical decision paradigm.



# Chapter 4

## Toddlers' processing of Dutch vowels: A mispronunciation paradigm

## Introduction

In previous chapters, we assessed phonological representations of the Dutch vowel system in adult perception of vowels /e, ø, o/. We investigated vowel contrasts of place and labiality. In the EEG experiment in Chapter 2, two perceptual asymmetries were revealed: a coronal-dorsal asymmetry, and a labiality asymmetry. The former is in line with previous literature; coronal-dorsal perceptual asymmetries have been attested in both consonants (e.g. Lahiri & Reetz, 2002) and vowels (e.g. Eulitz & Lahiri, 2004; De Jonge & Boersma, 2015) – see also Chapter 1 and 2. The asymmetry between labial and nonlabial vowels, however, has to our knowledge not been attested in adult native between-category discrimination before. As argued for in Chapter 1, we assume a separation of the feature labial from the features coronal and dorsal by introducing a LINGUAL node parenting [CORONAL] and [DORSAL], and a LABIAL node, parenting [LABIAL]. Based on the findings in Chapter 2, we proposed a role for these nodes in the geometry during mapping; dorsal and coronal SR features are mapped onto a feature in UR below the LINGUAL node, whereas a labial SR feature is compared to a feature below the LABIAL node. This is discussed in more detail in Chapter 2 and Chapter 5. More concretely, we proposed the following phonological representations for /e/, /ø/, and /o/ in Dutch ([Φ] indicates underspecification):

Surface representation (SR):

```
     /e/                 /ø/                     /o/
      |                   |                       |
     ART                 ART                     ART
      |                 /   \                   /   \
   LINGUAL         LABIAL   LINGUAL        LABIAL   LINGUAL
      |               |        |              |        |
  [CORONAL]       [LABIAL]  [CORONAL]      [LABIAL]  [DORSAL]
```

Underlying representation (UR):

```
     /e/                 /ø/                     /o/
      |                   |                       |
     ART                 ART                     ART
      |                 /   \                   /   \
   LINGUAL         LABIAL   LINGUAL        LABIAL   LINGUAL
      |               |        |              |        |
     [Φ]          [LABIAL]    [Φ]          [LABIAL]  [DORSAL]
```

In the current chapter, we will focus on children's phonological representations of these vowels. Investigating child phonology is worthwhile, because whereas adults have a complex and complete language system in place which may be hard to disentangle, children acquiring their mother tongue are still building their system. Young children may not have acquired all phonological contrasts yet, or their phonological representations may not be fully established yet. The child's phonology provides insight into how the adult phonology is built, and consequently can inform us on what is possible for the adult system. What do Dutch children's phonological representations of place and labiality look like?

In the current study, perceptual data from Dutch three-year-old toddlers is collected. We investigate perception during a word recognition experiment, using an intermodal preferential looking paradigm with place and labiality mispronunciations of /e, ø, o/. Thus, we investigate the same three contrasts as we did in previous chapters on adults:

(1) Single-feature contrast (place): coronal versus dorsal - /ø/ versus /o/;
(2) Single-feature contrast (labiality): labial versus nonlabial - /e/ versus /ø/;
(3) Two-feature contrast (place & labiality): /e/ versus /o/.

We investigate perceptual asymmetries in these contrasts as a window into phonological representations. As three-year-old Dutch children do use coronal, coronal labial and dorsal labial vowels in production (Beers, 1995), we can assume three-year-olds have acquired the same phonological representations to account for these contrasts as adults (SR & UR as presented above). Therefore, we expect them to show similar asymmetries. However, it is also possible that representations for coronal labial vowels like /ø/ might not be fully engrained yet, as they have only recently entered children's production at this age.

Before we discuss the experiment and the hypothesized outcomes in more detail, we will first turn to vowel perception over the course of first language acquisition.

**On the development of vowel perception**

Within the field of research on vowel perception, some distinctions are important to keep in mind: studies may investigate listeners' *native* perception (perception of native phones), or *non-native* perception (perception of phones from non-native languages). Both types of study can either test *within-category* or *between-category* vowel perception. Within-category studies investigate the sensitivity to vowel changes within one vowel category, e.g. different tokens of /i/. Between-category perception regards sensitivity to changes that cross category boundaries, i.e. sensitivity to changes from one phoneme to another; e.g. /o/ versus /u/. In this dissertation, we focus on native between-category perception of vowel quality, namely the perception of Dutch vowels that differ regarding place and/or labiality features.

Werker & Tees (1984) (consonants) and Kuhl and colleagues (1992) (vowels) illustrated that infants from different linguistic backgrounds start out with a comparable ability to discriminate a wide variety of human speech sounds, but attune their perception to the contrasts relevant to their native language during the first year of life. It has been shown that children become attuned to the vowels of their native language around six months of age (e.g. Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992; Werker & Tees, 1984; confirmed by meta-analysis by Tsuji & Cristiá, 2013). For instance, Kuhl and colleagues demonstrated that six-month-old Swedish and American English infants both show a strong magnet effect for native-language phonetic prototypes, but not for non-native ones.

Kuhl's Native Language Magnet model (NLM) (e.g. Kuhl, 1991, Kuhl et al. 2008) assumes that early in life listeners develop acoustic prototypes for native phonemic categories (e.g. Grieser & Kuhl, 1989, Iverson & Kuhl, 1996; Kuhl, 1991, 1992; Kuhl et al., 1992) based on frequency of occurrence. One of the driving forces behind infants' transition towards a language specific pattern of phonetic perception is statistical learning, i.e. the detection of distributional frequencies in the patterns of speech sounds in the speech one is exposed to. Exposure to native speech sounds leads to the formation of prototypes, which have been described as the representations most often activated (Kuhl et al., 2008), or as the centre of a vowel category (cf. Feldman, Griffiths, & Morgan, 2009). These prototypes serve as magnets and warp perceptual space; perceptual space is compressed around prototypical phonemes. As a result, within-category discrimination of sounds close to a prototype is more difficult than discrimination of speech sounds towards the edge: tokens close to a prototype are perceived to be more similar to the prototype. In sum, based on NLM, an asymmetry in perception is predicted, with poorer performance if the central instance is used as the standard in discrimination studies than if a more peripheral sound is used as standard. Evidence supporting such within-category perceptual asymmetries is provided by e.g. Gieser & Kuhl (1989), Kuhl (1991) and Kuhl et al. (1992).

The NLM model is designed for within-category discrimination. If both vowels are separate categories (between-category discrimination), predictions are less clear. However, since phonemes that are more frequent are assumed to be stronger magnets, one could hypothesize that frequent vowels allow for more variation. Based on this assumption, we would predict that a change from a frequent to an infrequent sound would be more difficult to perceive than vice versa. As illustrated in Chapter 1 based on counts in CELEX, tense front round vowels are relatively infrequent in Dutch compared to front (unround) vowels and back (round) vowels. Thus, based on the frequency distributions, a change from /e/ or /o/ towards [ø] would be expected to be less prominent than vice versa.

A model that captures perceptual directional asymmetries in both native and non-native between category contrasts is the Natural Referent Vowel (NRV) framework (Polka & Bohn, 2003; 2011, Masapollo, Polka, & Molnar, 2017, Masapollo, Polka, & Ménard, 2017). The NRV framework assumes innate initial perceptual biases based on the universal extremes: /i/, /u/ and /a/. These vowels serve as reference points to be used to acquire additional categories throughout experience with a particular language. As discussed in the dispersion theory by Lindblom and Engstrand (1989), the universal focal (cardinal) vowels /i/, /u/, and /a/ establish a minimal vowel system organized in such a way that elements are sufficiently distinct or *dispersed* from each other. According to Schwartz and colleagues (Schwartz, Boe, Vallee, & Abry, 1997; Schwartz, Abry, Boe, Ménard, & Valée, 2005) focal vowels provide global constraints, but are also universally favoured across vowel inventories, because of their acoustic properties. They have perceptual attributes that make them worthy phonetic elements on their own. Based on vowel perception experiments with adults, Schwartz and Escudier (1989) propose that focal vowels are more stable in short-term memory. Lieberman (1971) already noted that formant frequency patterns of focal vowels yield prominent spectral peaks that make it easier to perceive the sounds. In the Dispersion-Focalization theory (Schwartz et al., 1997; 2005), Schwartz and colleagues pose that formant convergence in focal vowels could result in increased so-called *perceptual value*, because of acoustic salience. In line with this literature, NRV assumes peripheral vowels (i.e. at the boundaries of vowel space) to act as natural referent vowels or perceptual anchors due to their focalized formants. More focalized vowels are acoustically regarded to be more salient than less focalized vowels, because closeness of formants increases acoustic energy in certain bandwidths.

Generally, a change from a more central to a more peripheral vowel is easier to detect than vice versa. This would also predict asymmetries for the native between-category contrasts of interest in the current chapter. Indeed, studies have revealed vowel asymmetries in early perception which are in line with NRV predictions. (e.g. Polka & Werker, 1994; Swoboda et al., 1978; Bohn & Polka, 2001; Swoboda, Kass, Morse & Leavitt, 1978 as reviewed in Polka & Bohn, 2003; 2011).

Since the current study focuses on native between-category contrast sensitivity, we provide an overview of native between-category discrimination in infant vowel perception data for place and labiality in Table 4.1. We found only two relevant studies: Polka and Bohn (1996) regarding place, and Polka and Bohn (2011) regarding labiality.

**Table 4.1** Overview of native between-category (a)symmetries in infant perception of vowels that differ regarding place and/or labiality.

| Contrast | Language | Age (months) | Task | Asymmetry | Reference |
|---|---|---|---|---|---|
| **Place** | German | 6-8 | CHP | /y/ → /u/ | Polka & Bohn (1996) |
| | | 10-12 | | /y/ → /u/ | |
| **Labiality** | Danish | 6-8 | CHP | /e/ → /ø/ | Polka & Bohn (2011) |
| | | 9-12 | | no asymmetry | |

Arrows indicate the direction of the change that was easier to detect. Both experiments used /dVt/ as word templates. CHP = conditioned headturn procedure.



In a conditioned head turn procedure, Polka and Bohn (1996) showed that German infants older than 6 months are more sensitive to a change from coronal /y/ to dorsal /u/ (place contrast) than vice versa in a conditioned head turn procedure. Interestingly, this asymmetry between /y/ and /u/ in German infants is in opposite direction compared to the coronal-dorsal asymmetry found for adult German listeners in a mismatch negativity study by Eulitz and Lahiri (2004) as well as for Dutch listeners as reported in Chapter 2 for /ø/ and /o/, where the MMN was enhanced when the deviant was coronal. Infants approaching their first birthday are becoming increasingly sensitive to phonetic differences that reflect the acoustic-phonetic properties of the native phonetic categories. However, their underlying speech processing skills are not yet organized with respect to higher levels of phonological organization, but are aligning perceptual biases with native phonetic categories (Polka & Bohn, 1996). As such, the difference between adult and infant perception could be a matter of developmental stage: infants may rely on phonetic categories, but have not completed their phonological organization yet.

Polka and Bohn (2011) report an asymmetry between /e/ and /ø/ in Danish infants around 6 months. For these infants, a change from /e/ to /ø/ is easier to detect than the reverse. The reported asymmetry for /e-ø/ in Danish young infants is thus in the same direction as we found for the labiality contrast between /e/ and /ø/ in Dutch adults in Chapter 2, where a change from /e/ to /ø/ yields an enhanced MMN compared to the reverse. Polka and Bohn initially predicted discrimination to be facilitated when the second vowel in a vowel pair is the more peripheral one regarding F1-F2. Discrimination of the more peripheral /e/ and the less peripheral /ø/ resulted in an order effect opposite to their prediction. This implies that the 'peripherality' in F1-F2 vowel space are not sufficient to predict directional asymmetries. Predictions can vary regarding whether one bases the predictions on F1-F2 or F2-F3. Based on F1-F2 /ø/ would be more focal than /e/. Based on the convergence or closeness of F2 and F3, /e/ is more focal than /ø/. Thus, the two dimensions make opposite predictions. NRV is not explicit on the relative weighting of F1-F2 focalization and F2-F3 focalization. Hence, predictions are not always clear.

Even though the asymmetry reported by Polka and Bohn (2011) is in similar direction as the labiality asymmetry attested in Chapter 2, we cannot simply assume the underlying mechanism to be the same for both participant groups, as adults may give perceptual priority to phonemic dimensions over other aspects of phonetic structure, whereas six-month-old infants - who have a limited vocabulary - may simply respond to acoustic-phonetic properties rather than employ phonological representations. This is supported by the fact that the asymmetry between /e/ and /ø/ in Danish infants seems to have disappeared in 12-month-old infants (Polka & Bohn, 2011).

Indeed, there is evidence that infants' sensitivity to phonemic contrasts does not necessarily reflect their phonological representations. During their first year of life, infants tune into their native sound system as they learn to focus on the sound contrasts important to their native language (Werker & Tees, 1999; Kuhl et al., 2006; Tsao, Liu & Kuhl, 2006), and how these sounds may combine into words (Jusczyk, Cutler, & Redanz, 1993; Jusczyk, Luce, & Charles-Luce, 1994), i.e. phonotactics. They also learn which stress pattern dominates their native language as well as other cues to detect word boundaries and extract words. One might assume that when they acquire their first words around their first birthday, they also naturally learn the sound structure of a word and store the phonological form in the mental lexicon.

However, nothing could be further from the truth. Perception studies have shown that learning speech sounds in words is more challenging than either sound discrimination or pattern recognition as required for phonotactic learning. For example, 14-month-old infants are able to discriminate [b]in and [d]in in a discrimination task, but fail to show this ability when they are required to pair the words with objects (e.g. process them at a lexical level – with a word-meaning association) (Stager & Werker, 1997; Pater, Stager & Werker, 2004). Whereas some claim that the difference in sensitivity in lexical versus non-lexical processing is due to task demands (e.g. Werker, Fennel, Corcoran & Stager, 2002), others argue that this is due to a disparity between phonetic (non-lexical) and phonological (lexical) representations (e.g. Fikkert, 2010).

Werker and colleagues assume that a task demand that is too high can lead to poor performance, because infants' limited processing capacities prevent them from listening carefully to the shape of speech sounds when they have to form word-object associations. A study showing that children can distinguish between the same sounds in known words rather than novel words seems to support this notion (Fennell & Werker, 2003; Werker & Fennell, 2004). However, as the authors note, these results can also generate an alternative interpretation, that phonological representations of stored words contain less detail and are more abstract than the phonetic representations of unanalysed strings of the pre-lexical child (Pater et al., 2004).

The assumption that infants' or toddlers' lexical phonological representations are not adult-like yet finds support in the literature. Stager and Werker (1997) did not take into account whether perception of the contrast between /b/ and /d/ was asymmetrical. For discrimination of bin and din, an asymmetry has been reported for Dutch 14-month-old infants in Fikkert (2010). Asymmetries in the early lexicon can be explained assuming the logic of the Featurally Undespecified Lexicon (FUL) model as used in Chapter 2: a specified feature does not accept a mispronunciation, so a change is readily detected, whereas an underspecified feature does accept a mispronunciation, and that change is not salient to the child. Some researchers believe that since at early stages in development a child's lexicon is limited in size, it contains few phonological neighbours and detailed representations are not needed to distinguish different lexical items (e.g. Charles-Luce, & Luce, 1990). Word recognition is thought to begin around six months of age (e.g. Tincoff & Jusczyk, 1999; Bergelson & Swingley, 2012). When vocabulary grows, more detail in phonological lexical representations is required to differentiate between lexical entries. Rather than using vague or holistic representations (e.g. Charles-Luce, & Luce, 1990; 1995; Coady, & Aslin, 2003), however, Fikkert and Levelt (2008) argue that children do use phonological representations from the start of their early lexicon. They do not, however, store every phonological feature in their lexical representations from the start.



For discrimination (non-lexical), short term memory traces are sufficient, whereas at a lexical level features need to be stored in the mental lexicon, and items need to be retrieved through phonology. Ensuring phonological processing, the current study investigates toddlers' perception of vowel contrasts at a lexical level in a mispronunciation paradigm in toddlers.

**Mispronunciation paradigms**

Testing the impact of mispronunciations on word recognition can inform us about the nature of the phonological lexical representations as we also discussed in Chapter 3. Rather than using a semantic priming experiment as we performed with adults, an intermodal preferential looking paradigm is a more suitable approach in toddlers. Swingley & Aslin (2000) used this method to reveal a mispronunciation effect: word recognition was hindered when words were mispronounced (MP), shown in proportion of target looking (PTL) as a faster response and a larger proportion of target looking in response to correctly pronounced (CP) trials than to MP trials.

Children show sensitivity to mispronunciations of words by 14 months regarding consonants, tone, and vowel quality and duration (e.g. Bailey & Plunkett, 2002; Ballem & Plunkett, 2005; Fennell & Werker, 2003; Swingley & Aslin, 2000, 2002). For example, Ballem and Plunkett (2005) report 14-month-olds to be sensitive to mispronunciations of word initial consonants in both familiar and novel words.

However, in this early literature using the intermodal preferential looking paradigm, phonological factors are not controlled, nor are directional asymmetries tested. For example, using a mispronunciation paradigm, Mani and Plunket (2008) tested discrimination of two contrasts: /æ/-/u/ and /ɔ/-/ɪ/ in British English 14- and 18-month-olds. Discrimination is only tested in a single direction. Detection of these mispronunciations generally tells us that 14-month-olds are sensitive to mispronunciations of vowels, but it does not inform us about the exact phonological representations at stake.

In fact, some researchers have argued that sensitivity to mispronunciations is guided by the specific phonological contrast or change. Assessing the impact of mispronunciations on word recognition for both directions of a contrast can provide more insight into phonological representations, since it allows us to assess perceptual asymmetries. Using intermodal preferential looking paradigm, Van der Feest and Fikkert (2015), were the first to investigate how different directions of a contrast impact word recognition in young children. They found a labial-coronal asymmetry between word-initial /t/ and /p/ in 20-month-old and 24-month-old Dutch children: When */t/and* (tooth) was mispronounced as *[p]and,* this did not hinder word access. Thus, labial [p] was accepted as a variant of coronal /t/. When */p/oes* (cat) was mispronounced as **[t]oes,* word recognition was hindered, shown by a significant lower proportion of target looking for this mispronunciation than for the correct pronunciation. Thus, a coronal mispronunciation is not accepted as a variant of a labial sound. Also, using the same paradigm, 24-month-old children showed an asymmetry with respect to voicing: a change from unvoiced to voiced stop hindered word recognition, but not vice versa (Van der Feest & Fikkert, 2015). Following this example, others have used similar paradigms assessing directional asymmetries in infants and toddlers (e.g. Tsuji, Fikkert, Yamane, & Mazuka, 2016; Ramachers, Brouwer, & Fikkert, 2017), in both well-known words as in newly learned words.

At present, no lexical perception data contrasting coronal labial vowels with either coronal (nonlabial) or dorsal (labial) vowels is available in the L1 acquisition literature, and so the current study is the first to provide this. In a word-learning habituation task, Curtin, Fennell and Escudero (2009) showed that Canadian 15-month-olds fail to discriminate /i/ and /u/ (*deet* vs. *doot*) – a contrast between a coronal vowel and a dorsal labial vowel (similar to /e/ versus /o/ in the current study). However, it is unclear whether toddlers were not sensitive to the contrast in both directions, since the authors do not report on sensitivity in different directions. It is unclear how to interpret their results with respect to phonological representations of these toddlers. Testing a contrast between a front unround vowel and a back round vowel in a habituation paradigm, Fikkert (2010) reports an asymmetry in discrimination of /ɪ/ vs. /ɔ/ in the syllables *bin-bon* in 14-month-old Dutch listeners: infants were sensitive to a change from *bon* to *bin*, but not vice versa. This suggests place of articulation of /ɔ/ to be specified and /ɪ/ to be underspecified. Compliant with the representations argued for by Fikkert and Levelt (2008), Fikkert (2010) argues that at this age infants use global representations: the place of articulation of the stressed vowel is used to represent the whole word.

**Present study**

***Word learning IPLP***

In the current study, we use a word learning adaptation of the traditional Intermodal Preferential Looking Paradigm (IPLP) (e.g. Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987; Bailey & Plunkett, 2002; Swingley & Aslin, 2000; Golinkoff, Ma, Song, & Hirsh-Pasek, 2013) – also called the looking-while-listening procedure (e.g. Fernald, Zangl, Portillo & Marchman, 2008; Della Luche, Durrant, Poltrock, & Floccia, 2015) in 38-month-old Dutch toddlers. The IPLP is a valid, sensitive measure of young children's language knowledge (Golinkoff et al., 2013), and has been successfully used in testing word learning or word recognition experiments, for example when testing segmental or suprasegmental (e.g. tone contrasts) phonology in mispronunciation paradigms (e.g. Mani & Plunkett, 2008; Altvater-Mackensen, et al., 2014; Van der Feest & Fikkert, 2015; Ramachers, et al., 2017; Tsuji et al., 2016). Although IPLP is most commonly used between the ages of 15-24 months, it also proved successful in 2.5 – 4-year-olds in Ramachers et al. (2017), who studied sensitivity to lexical tone in Dutch toddlers.



To test phonological contrasts using IPLP, stimuli are required to meet several criteria, such as: (1) Target word should be a monosyllabic word with /CVC/ template; (2) Different target words are of similar frequency/familiarity to the child; (3) Target word has a depictable referent; (4) Mispronunciation of the target word results in a pseudoword rather than a real word; (5) For each target vowel, two suitable CP target words are required, to reduce the possibility that any effects were idiosyncratic to a particular word. In order to meet all these criteria, the current experiment used novel words. Rather than using known words, the intermodal preferential looking paradigm can also be used with novel words, i.e. a word learning variant of IPLP (Delle Luche, Durant, Poltroch, & Floccia, 2015). IPLP has been used in novel word learning before (e.g., Mani & Plunkett , 2008; Schafer & Plunkett, 1998;  Swingley & Aslin, 2002, 2007; Tsuji et al., 2016). The procedure with a novel target and a novel distractor object has also successfully been applied in similar word learning studies with three- to five-year-olds (Singh & Quam, 2016). In particular, the procedure was proved successful in 2.5 to 4-year-old Dutch toddlers by Ramachers et al. (2017). Our experiment is based on the paradigm used by Ramachers et al. (2017) and Tsuji et al. (2016).

Compared to using known words, using novel words increases sensitivity of the method for measuring differences in the proportion of target looking for mispronunciations versus correct pronunciations. For example, when known words are used, in a paradigm with one target and one distractor on screen, mispronunciation of the target will probably still lead to a naming effect, because the mispronounced target word is still more like the CP than like the label for the distractor item. For example, in the experiment by Van der Feest and Fikkert (2015), distractor item *pop* (/pɔp/ - 'doll') was used. When *poes* (/pus/ - 'cat') is mispronounced as *toes (/tus/), the mispronunciation is still more like *poes* than like *pop*. Significant naming effects were present in this case. However, when using novel words, the distractor item does not need to have a label since it can be a novel object. As a result, if the distractor has no label, a mispronunciation of the target may more easily be interpreted as a potential label for the distractor object. If a mispronunciation hinders word recognition, the distractor object may be considered as a target. Consequently, it is less likely that a positive naming effect is found for the target object. Rather, it is more likely to find negative naming effects: longer looking times towards the distractor than towards the target.

***Research questions and predicted outcomes***

In the current study, we are investigating lexical access for vowel mispronunciations regarding labiality and place in tense mid vowels /e/, /ø/, and /o/. The same three research questions are of central focus as in the other chapters of this thesis:

RQ1. How is the vowel contrast between coronal and dorsal (i.e. place) represented?

RQ2. How is the vowel contrast between labial and nonlabial (i.e. labiality) represented?

RQ3. How do place and labiality in vowels interact when both are contrastive?

We will first measure naming effects: does looking towards the target object increase after naming compared to pre-naming? If so, this indicates word recognition, and would show that the procedure works. Is this effect also present for the different mispronunciations, or do mispronunciations hinder word access? We assess asymmetries in these naming effects for both MPs of each contrast.

Production data supports the view that, as early as at 14 months, Dutch children produce the contrast between coronal and labial vowels (Beers, 1995; Fikkert & Levelt, 2008;). This is also supported in perception experiments (e.g. Fikkert, 2010; Van der Feest & Fikkert, 2015). Thus, they should be sensitive to the phonological contrast between /e/ and /o/ early on, since this contrast can be perceived based on a contrast between [CORONAL] and [LABIAL]. At 14 months, however, children have not learned yet that coronal labial vowels are part of the phoneme inventory too (hence, contrastive with /e/ and /o/). Coronal labial vowels signal that coronal and labial are not mutually exclusive features for Dutch vowels, so when coronal labial vowels enter a child's phonology, a coronal-labial contrast (or a labial-lingual contrast) is not sufficient anymore. A three-way distinction (e.g. /e, ø, o/) requires at least two phonological contrasts. Beers (1995) reports children acquiring Dutch to use the first coronal labial vowels in production around the age of 3;0 – 3;2 (36-38 months). Presumably, children from this age onwards should be able to perceptually distinguish coronal labial vowels (like /y/ and /ø/) from coronal (/i/ or /e/), and dorsal labial vowels (/u/ or /o/) at the lexical level. We can expect the place contrast (coronal vs. dorsal) and the labiality contrast (labial versus nonlabial) in vowels to be present in the child's phonology around this age to account for the three-way contrast. Therefore, we predict that at 38 months of age, Dutch toddlers have adult-like representations accounting for the contrasts of interest. We predict asymmetries for both place and lability, similar to adult outcomes in Chapter 2. Predicted outcomes are presented in Table 4.2.

For place (RQ1: vowel pair /ø-o/), we predict an asymmetry. We predict that when dorsal /o/ is mispronounced as coronal [ø], this results in a mismatch, hindering word recognition. Hence, we predict a smaller or absent naming effect for this MP. Vice versa, when coronal /ø/ is mispronounced as dorsal [o], we predict mapping of SR and UR to result in a no-mismatch, because the coronal feature of /ø/ is underspecified. We do not expect this mispronunciation to hinder word recognition. Instead, a naming effect similar to CP is predicted.



For the labiality contrast between /e/ and /ø/ (RQ2), we also predict an asymmetry. When a nonlabial vowel /e/ is mispronounced as labial [ø], we predict that this will result in an aborted mapping, because the surface representation of [ø] has a LABIAL node, which is absent in UR of /e/. Since we assume this to be interpreted as a phonological conflict in perception (see Chapter 2), we predict this to hinder word recognition, and hence show a smaller naming effect than expected, if at all. If this conflict is so strong that the MP is considered beyond doubt to be a different word than the CP, this condition may result in a negative naming effect. This would indicate that the MP is interpreted as a label associated with the distractor object, shown by a looking preference towards the distractor. Vice versa, when labial /ø/ is mispronounced as [e], word recognition will not be hindered. [e] is nonlabial, so mapping of labiality is not initiated. Thus, no conflict is perceived when /ø/ is mispronounced as [e], and word recognition is not hindered. A positive naming effect is predicted.

In sum, we predict asymmetries for the two single-feature contrasts of interest.

For the two-feature contrast between /e/ and /o/ (RQ3), we predict that both directions of mispronunciations will result in a phonological conflict. When /e/ is mispronounced as [o], mapping is aborted, because mapping of [LABIAL] SR of [o] requires a LABIAL node in UR of /e/ which is absent. Vice versa, the [CORONAL] SR feature of [e] mismatches with [DORSAL] feature in UR of /o/. Thus, we predict hindered word recognition for MPs in both directions of change for this contrast. This would be a symmetrical result. Furthermore, White and Morgan (2008) showed in their preferential looking experiment that 19-month-old toddlers can show graded sensitivity to degree of MPs: two-feature contrasts in word-initial consonants may show larger differences between CPs and MPs than single-feature changes. Thus, for the two-feature contrast, we may find larger differences between CPs and MPs.

**Table 4.2** Predicted outcomes of IPLP word learning MP paradigm.

| Vowel pair | Contrast | Featural Distance | Condition | CP vowel | MP vowel | SR of MP | UR of CP | Mapping prediction | Naming effect | Asymmetry predicted |
|---|---|---|---|---|---|---|---|---|---|---|
| **/ø/- /o/** | place | 1 | [o]/ø/ | /ø/ | [o] | [DORSAL] | [Φ] | no-mismatch | yes | YES |
| | | | [ø]/o/ | /o/ | [ø] | [CORONAL] | [DORSAL] | mismatch | no/negative | |
| **/e/-/ø/** | labiality | 1 | [ø]/e/ | /e/ | [ø] | [LABIAL] | | *aborted* | no/negative | YES |
| | | | [e]/ø/ | /ø/ | [e] | | [LABIAL] | | yes | |
| **/o/-/e/** | place & | 2 | [e]/o/ | /o/ | [e] | [CORONAL] | [DORSAL] | mismatch | no/negative | NO |
| | labiality | | | | | | [LABIAL] | - | | |
| | | | [o]/e/ | /e/ | [o] | [DORSAL] | [Φ] | no-mismatch | no/negative | |
| | | | | | | [LABIAL] | | *aborted* | | |

CP = correct pronunciation; MP = mispronunciation; SR = surface representation; UR = underlying representation.

If aborted mapping (based on a labiality conflict when /e/ is mispronounced as [o]) is a stronger conflict than a mismatch (based on a place conflict when /o/ is mispronounced as [e]), or vice versa, results of the two-feature contrast could be asymmetrical: one direction may result in a negative naming effect, whereas the other direction of change yields no (or possibly a small) naming effect.

## Method

In the current section, we will describe the mispronunciation experiment using a word learning variant of the intermodal preferential looking paradigm in more detail.



### Subjects

Participants were recruited from the subject pool of the Baby and Child Research Center (BRC) at Radboud University in Nijmegen, and were also tested at the BRC in Nijmegen. Ethical approval for the study was obtained from the Ethiek Commissie Faculteit der Sociale Wetenschappen (ECSW) at Radboud University.

137 typically developing monolingual native Dutch toddlers aged 37-39-months were included in the final sample (mean age = 1161.5 days; range = 1131 – 1187 days; SD = 13 days; 70 female). For all subjects, parents/caregivers reported 100% use of Dutch in the home environment. All children had normal hearing, and normal or corrected-to-normal vision. None of the participants had severe and recurrent, or recent (< 4 weeks before testing) otitis media. None of the participants were born over two weeks early. Children at familial risk of developmental language disorders (DLD) or dyslexia (defined by a sibling or parent diagnosed with DLD or dyslexia) were excluded from the study, as were children who received speech therapy. None of the subjects had known cognitive or neurological impairments or developmental disorders or delays.

In total, 151 subjects participated in this intermodal preferential looking experiment. Fourteen subjects were excluded due to: receiving speech therapy (reported at test) (3 subjects); reduced vision (reported at test) (1 subject); hearing loss due to acute rhinitis (1 subject); use of a self-conceived label for the novel objects (1 subject); interference of caregiver/parent during testing (1 subject); too few good test trials (7 subjects). For a detailed description of the trial, block and participant exclusion criteria used, refer to the section Data Preprocessing.

Parents/caregivers signed informed consent prior to participating in the study. They received a picture book or small monetary contribution as compensation for their partaking.

**Apparatus**

All children were tested in a dimly lit and quiet room. They sat on their caregiver's lap on a chair placed centrally in front of the 17-inch television screen (LG 47LK530 ZC) on which the experiment was presented (both audio and visual stimuli), with volume of the television set at 60. The television screen was surrounded by black wooden partitions and black curtains. The experimenter stood behind this black screen while running the experiment, so the child would not be distracted by the experimenter or the experimental computer. A digital camera was placed 30 cm below the television screen, hidden behind a black curtain with an opening for the lens, to record the child's behaviour. Images captured by the camera projected real-time on a computer screen for the experimenter to monitor the child during the experiment. The experiment was run using Lincoln Infant Lab Package 1.0 (Meints & Woodford, 2008) – also called LOOK. The caregiver wore over-ear Sennheiser headphones and listened to masking music during the experiment. They were instructed to not engage in any interaction with their child during testing.

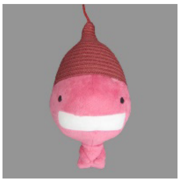
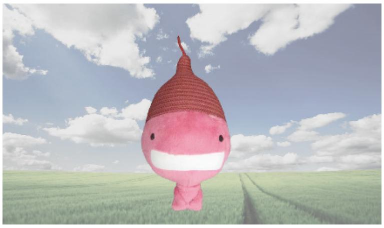
**Procedure**

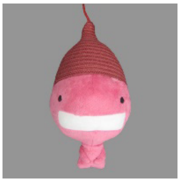
The experiment was an intermodal preferential looking paradigm (IPLP) (Golinkoff et al., 1987; Golinkoff et al., 2013) with a word learning phase. The paradigm was adopted from Ramachers et al. (2017), who took the paradigm by Tsuji et al. (2016) as her starting point.

Three vowel contrasts were tested: /ø-o/ (place contrast), and /ø-e/ (labiality contrast), /e-o/ (two-feature contrast, where both place and labiality are contrastive). Each contrast has two *directions of change*, e.g. a target word with /o/ in the correct pronunciation (CP) can be mispronounced (MP) by substituting the /o/ by [e], or vice versa, when CP /e/ is mispronounced as [o]. Each child was tested on one of the three contrasts in both directions of change (within-subject design). Each child was tested on two blocks, each testing opposite directions of change. An overview of one experimental block is given in Table 4.3. In each block a child learned a novel word-object mapping during a *learning phase*. Subsequently, their word recognition was measured during the *test phase*. Which direction of the contrast to be presented in the first or the second block was counterbalanced across subjects. Blocks were separated by a one-minute break. The entire experiment took 10 to 15 minutes.

Each trial was manually initiated by the experimenter only when the child looked at the television screen. In between trials, a blinking light was presented at the screen centre as an attention-getter. Each trial had a set duration (see Table 4.3). During each trial, eye gaze was coded by the experimenter (see 4.2.5. Data preprocessing for more detail).

**Table 4.3** Overview of a single block.

| Phase | No. of trials | Trial duration (ms) | Example of visual stimuli | Example of auditory stimuli |
|---|---|---|---|---|
| **Introduction** | 1 | 4800 |  | **In 1st block**: "Hey, hallo! Laten we een spelletje spelen. Let goed op!"<br>*"Hey, hello! Would you like to play a game? Pay attention!"*<br>**In 2nd block**: "Goed gedaan! Nog een keer!"<br>*"Well done! Here we go again!"* |
| **Object Familiarization** | 1 | 9000 |   | "Kijk eens hier! Wat zijn dat? Die zijn mooi! Vind jij ze ook leuk?"<br>*"Look at this! What are those? Those look nice! Do you like them too?"* |
| **Learning** | 4 | 30000 |  | **2 x Target**: "Kijk! Dit is een *tees*. Een *tees*! Zie je hem? Daar is de *tees!*... etc. (target word named 10x)<br>*"Look! This is a tees. A tees! Do you see it? There is the tees!... etc."* |
| | | |  | **2 x Distractor**: "Ooo wauw! Kijk! Die ziet er leuk uit. Wat zou dat zijn? … etc.<br>*"Oh wow! Look! That one looks nice! What would it be?... etc. "* |

| | | | | | |
|---|---|---|---|---|---|
| **Test** | 8 | 7000 |  |  | **2 x CP**: e.g. "Kijk naar de *tees*! Kun je hem vinden?"<br>*"Look at the tees. Can you find it?"* |
| | | | | | **2 x MP**: e.g. "Kijk naar de *teus*! Leuk hè?"<br>*"Look at the teus. Nice, huh?"* |
| | | |  |  | **4 x Filler**, e.g.: "Waar is de auto? Zie je hem?"<br>*"Where is the car? Do you see it?* |

Each child was tested on two blocks with similar structure. As an example, target word *tees* and its mispronunciation *teus* are used in this overview. CP refers to correct pronunciation. MP refers to mispronunciation. Table adapted from Ramachers et al. (2017).

Each block started with an ***introduction*** featuring a waving girl encouraging the child to play a game and pay attention. Then, during ***object familiarization***, the child was familiarized with two novel toy objects appearing simultaneously on the far left and far right side of the television screen. Familiarization before labelling usually boosts levels of retention (e.g. Hilton & Westermann, 2016). The objects were presented for 9 seconds. Neither of the objects is named yet. The child heard (in Dutch) *"Look! What are those? Those look nice! Do you like them too?"*. One of the objects is labelled in the subsequent learning phase (the target), and the other object does not receive a label (the distractor). Target side during familiarization was counterbalanced across blocks; the target of the second block was presented during familiarization on the opposite side compared to the first block.

After familiarization, the child was taught a novel word-object mapping during the ***learning phase.*** There were four 30 second learning trials. The first and third trial featured the target object, and the second and fourth trial featured the distractor object.

During a target learning trial, the target object was bouncing up and down in front of a natural scene, while it is labelled 10 times in sentences like *"Look at the [target]! This is a [target]! Can you see it? A [target]!"* Etcetera. We used the same carrier sentences as used in Ramachers et al. (2017), which were short declarative sentences with the target label in focus-final position.

The target learning trial featured twice during the learning phase, so in total the child heard the target object label 20 times. This repetitive labelling is in line with research on retention of novel word-object mappings (e.g. Quam & Swingley, 2010; Singh, Hui, Chan & Golinkoff, 2014; Hilton & Westermann, 2016), allowing children to make a word-referent association similar to e.g. Houston-Price, Plunkett and Harris (2005), Ma, Golinkoff, Houston, and Hirsh-Pasek (2011), and Ramachers et al. (2017).

Target learning trials were alternated with distractor trials (learning trials two and four), which were included to prevent novelty preference effects during the test phase. In those learning trials for distractors the object on screen was not labelled. Rather, the child heard short child directed sentences like *"Wow, do you see that? What would that be?"*, encouraging the child to pay attention to the object and wonder about its label.



After the learning phase, the ***test phase*** followed. This phase consisted of eight trials. There were four test trials and four filler trials. In each trial two objects were presented side by side and one of them was labelled. The two objects were presented on screen on the far left and far right, facilitating online coding of eye gaze.

In test trials, the target and distractor object from the familiarization phase are presented again, and the target object is named by a correctly pronounced (CP) or mispronounced (MP) label in a carrier sentence, e.g. *"Look at the [target]! Can you find it?"*. The MP involved a change in vowel quality, regarding place, or labiality, or place *and* labiality.

The presence of a non-labelled distractor object enables the child to consider this mispronunciation to be a label for the distractor item. This presupposes the use of the principle of mutual exclusivity (Markman, 1990), which guides children to map novel words to unfamiliar rather than familiar referents. The use of this principle has been demonstrated from the age of 16 months (e.g. Halberda, 2003; Diesendruck & Markson, 2001; Durrant 2014; Kalashnikova, Mattock & Monaghan, 2015). Using novel objects makes the procedure more sensitive to finding differences between CPs and MPs.

Test trials are alternated with filler trials. Filler trials involved correct pronunciations of well-known words and their referents *car, horse, cow* or *ball* in sentences like "*Where is the [word]? Do you see it?",* similar to Buckler & Fikkert (2016), Singh et al. (2014), and Ramachers et al. (2017). Filler trials are included to keep the child engaged and motivated, and to assess whether children perform the task and look at named objects if they recognize the label.

After block 1, a one-minute ***break*** followed, during which a short clip of farm animals with animal noises kept the child entertained.

Next, block 2 begins. Block 2 had the same structure as block 1, featuring a different pair of novel objects and a different novel word. Block 2 always tested the opposite direction of the vowel change tested in block 1, so each child provides data for both directions of the same contrast.

Twelve different versions of the experiment were used, testing three vowel contrasts with four experiment versions each. We counterbalanced which direction of change was tested in block 1 or in block 2. We also varied the order of CP and MP trials during the test phase. Test phases across all versions started with a filler trial, so the nature of the task would be clear to the child before the first test trial occurred. The first test trial (following the first filler trial) was always CP, but the order of the remaining three test trials varied in different versions of the experiment. Three pseudo-random orders of CP and MP test trials were used. In all test trials, the same two objects were presented on screen during the familiarization and learning phase, albeit at unpredictable sides of the screen. Test trials were pseudo-randomized in such a way that the target would never appear on the same side of the screen more than twice in a row.

Block 1 and block 2 always used a different word template: /tVs/ or /dVn/. Which novel word mapped onto which of the four novel plush objects (depicted in Figure 4.2 in the section Visual stimuli.) was counterbalanced over different versions of the experiment – so all objects served as targets and distractors, and as referents for all different novel words. Each pair of objects only occurred in one of the two blocks within one experiment. Which pair occurred in block 1 and which in block 2 was counterbalanced across versions.

For each of the twelve versions, data of at least ten participants were included in the final sample.

**Stimuli**

We used two templates: /tVs/, and /dVn/, for all target vowels. Since words are novel (pseudowords), they required novel objects as referents. This enabled us to match the objects in terms of attractiveness and animacy, and ensure visual distinctiveness in shape and colour. Using novel objects further increases the probability for the child that a mispronunciation may refer to the distractor object, reducing the chance of a positive naming effect in case of a phonological conflict. Also, using novel words requires a word learning phase before word recognition can be tested. This increases cognitive load, reducing the chance that the task is too easy for the participants.

Below, we provide more detail about both the auditory and visual stimuli, starting with the auditory stimuli.

***Auditory stimuli***

All auditory stimuli were recorded as child directed speech produced by a female native Dutch speaker, who grew up in the vicinity of Nijmegen and was still living in Nijmegen. To create all sentences for all trials, we used the same sentences as used by Ramachers et al. (2017) (see Table 4.3). Speech recordings were performed in a sound attenuated recording studio using Adobe Audition (version CS6). Recordings were judged and approved with respect to pronunciation and degree of child-directedness by three native Dutch speakers [NR,TS,PF].

Critical words were embedded within carrier sentences. For the learning trials, sentences for example were: *"Look! This is a [target]. The [target]. Do you see it? Look at the [target]... (goes on to include 10 namings of the target)"* For the test phase, the sentence in test trials was: *"Look at the [target]. Do you see it?"* or *'Look at the [target]. Nice, huh?"*. Filler trials were: *"Do you see the [target]? Do you like it?"*, or *"Where is the [target]? Can you find it?"*.



For test trials, CP and MP versions of the learned novel word were used as targets. Filler trials had correct pronunciations of known words car, ball, horse, and cow. These are words generally known to three-year-old children according to the Lexilijst Nederlands (Schlichting & Lutje Spelberg, 2002). In the test phase, target onset was always at 2500 ms – both for test trials and for filler trials.

The novel words used in the experiment contained one of three vowels /e/, /o/, or /ø/. Each novel word had a CVC template, where the consonants were fixed: /dVn/ and /tVs/. This resulted in six words that featured as either CP or MP words: /den/, /don/, /døn/, /tes/, /tos/, and /tøs/ depending on the version of the experiment.

Each of these novel words was either a pseudoword in Dutch, or a low frequency word that three-year-olds are unfamiliar with. One target word – */tos/* – is an infrequently used Dutch girl's name. Children who were familiar with that name (based on parental report) participated in versions of the experiment where this particular word did not occur, i.e. they were tested on a version of the /ø-e/ contrast instead.

By means of using several different tokens of the critical words we introduced acoustic variation to encourage more abstract phonological word learning and representation rather than storage of a particular acoustic form of a single token. In total, twelve tokens of each of the words /den/, /don/, /døn/, /tes/, /tos/, and /tøs/ were selected by a native speaker of Dutch [NR]. The words had to be clearly intelligible and spoken at a suitable speed. Of those twelve tokens for each word, ten featured in the learning phase and two featured in the test phase. Within one learning phase, both target learning trials were identical, using the same ten tokens of the target word. In the subsequent test phase, different tokens were used in both CP trials.

Selected tokens of the critical words were spliced into their carrier sentences, so for different conditions/versions of the experiment only the critical words varied, but the carrier sentences were identical. Stimuli were equalized for intensity to 65 dB (script by Kerkhoff, 2009) and prepared for the experiment using PRAAT (version 5.3.35; Boersma and Weenink, 2012). An example of an audio file of a test trial is given as a spectrum in Figure 4.1.

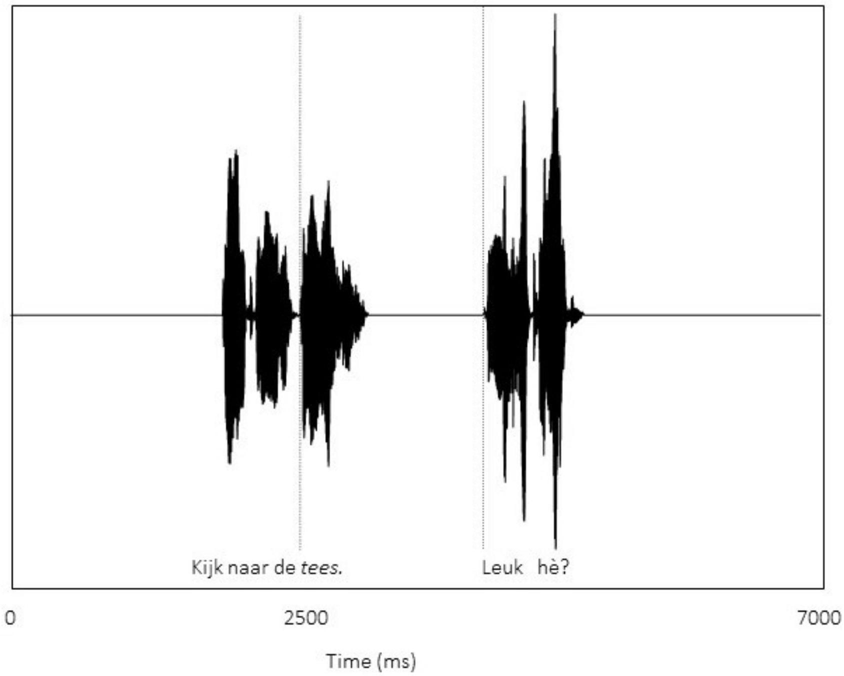



**Figure 4.1** Spectrum of a test trial. Critical word onset at 2500 ms into the trial. In between the two sentences is a silence of 1000 seconds. The entire test trial had a duration of 7000 ms, so silence is added at the start and the end of the audio.

***Visual stimuli***

We used four photos of animate plush puppets as visual target and distractor objects (see Figure 4.2). All visual stimuli were adopted from Ramachers et al. (2017). All objects had distinct shapes and colours (blue, pink, yellow and purple). Since shades for all objects differed, they were also distinct for colour blind people. Objects appeared in fixed pairs. The pink and blue object were paired together, as were the purple and yellow object. Pairs were matched in visual complexity, brightness and size (Ramachers et al., 2017). In Ramachers et al. (2017), it was confirmed that Dutch subjects of a similar age as the current study did not show a preference for a target objects within each pair of objects.

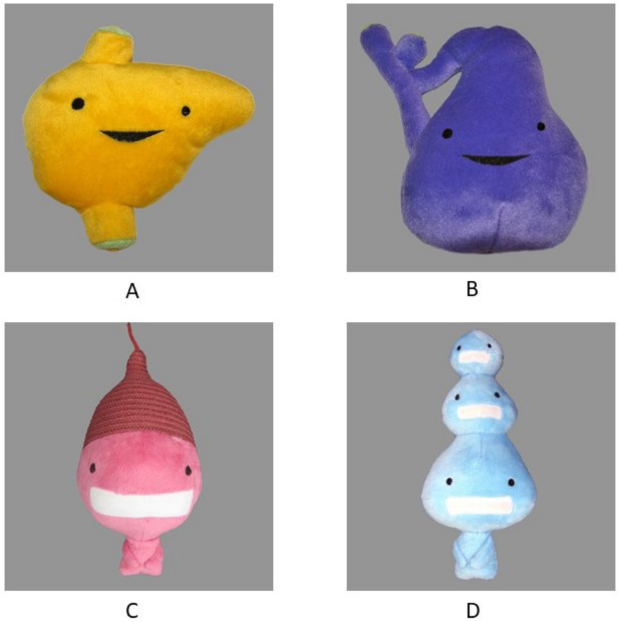



**Figure 4.2** Novel objects used as targets and distractors. Within one block, objects A and B were always paired with each other, and objects C and D were always paired with each other.

In filler trials (test phase), photos of familiar objects were used along with familiar words. Visual stimuli were photographs of two object pairs: a cow and a horse, and a ball and a car (see Appendix 6). To minimize boredom effects, different pictures per filler object were used across blocks.

During the object familiarization phase and the test phase, visual stimuli had a grey background – both for test trials and filler trials. In the learning trials, the plush objects were presented bouncing up and down against a background of a natural scene.

**Data preprocessing**

Looking behaviour was manually coded by the experimenter during testing (online), using Lincoln Infant Lab Package 1.0 (Meints & Woodford, 2008; also called LOOK). The experimenter was blind to the target side. The experimenter coded looking behaviour as ‘left’ or ‘right’ when two objects were presented side by side (i.e. familiarization and test trials). When the child was not attending the screen, this was coded as ‘other’ look. When the child attended the screen during learning trials, this was coded as ‘centre’ look, because in such trials the object was presented at screen centre rather than left or right. Trials other than familiarization, learning and test trials were irrelevant to our research questions and were disregarded.

***Setting time windows***

We measured naming effects in test trials, which are a change in proportion of target looking (PTL) after an object is named (post-naming window) compared to before it was named (pre-naming window). The pre-naming window was set to run from 0 to 2500 ms, since target onset was at 2500 ms.

For setting a meaningful lower end of the post-naming window we had to take into account three factors:

(1) minimal time required to initiate eye movements;
(2) reaction time of the experimenter during coding of eye gaze;
(3) display delays from camera to experimenter's computer screen.

Previous literature assumes young children's eye movement initiation takes 367 ms (e.g. Ramachers et al., 2017; Altvater-Mackensen et al., 2014; Quam & Swingley, 2010, Swingley & Aslin, 2000). However, whether this exact duration holds for the current age group is unclear, since this timing is based on younger children. The three-year-olds in the current experiment could be faster, closer to the 200 ms generally assumed for adults' saccade initiation (Altmann & Kamide, 2004; Dahan, Magnuson, Tanenhaus, & Hogan, 2001; Huettig & Altmann, 2005).

Since also the other two factors listed above were rather unclear, we determined our post-naming window based on the plot in Figure 4.3. This plot shows the proportion of target looks in filler trials, CP trials and MP trials. 40 ms time bins were used to determine the proportion of target looking over time, only including looks on screen (i.e. fixation on target or on distractor) – see also the formula for PTL presented in the Analysis section of the current chapter. Plotted data is not split into the different vowel conditions, so setting the time window is not influenced by impressions of what the results would be with respect to our particular research questions.

Figure 4.3 shows the general time course of word recognition, which appears similar for CP and filler trials. It shows at which time point CP and MP responses deviate. Starting at 3000 ms, PTL increases for all trial types, so from this time point onwards there seems to be an impact of the target word. Based on the three factors mentioned above it was unlikely for the target word to have an effect earlier than at 3000 ms. A window not starting before 3000 ms ensures that the analyses only consider changes in the child's looking behaviour that can reasonably constitute a response to the spoken word.

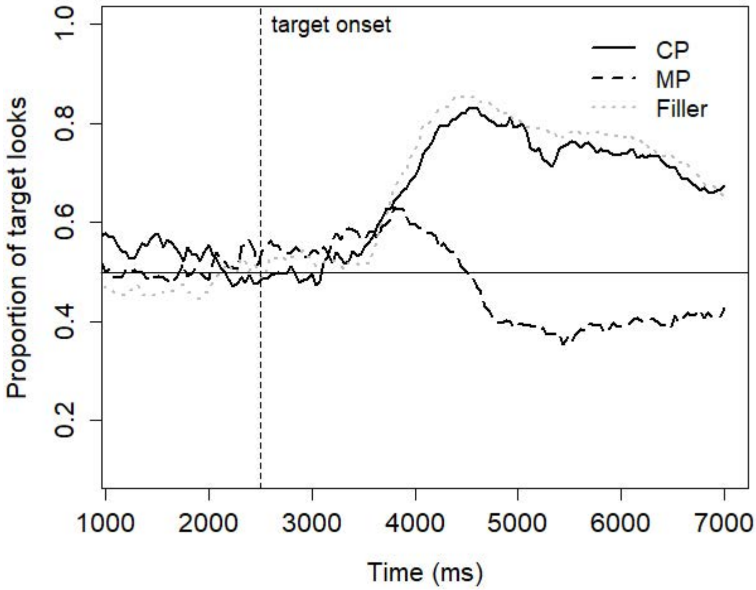



**Figure 4.3** Plot of data to determine window of interest. Horizontal line at 0.5 indicates chance level. Target onset at 2500 ms is indicated by the dotted vertical line. Proportion of target looks (PTL) for different trial types: fillers (known words like 'car' or 'ball'), correctly pronounced learned words (CP), and mispronounced versions of the learned words (MP). The proportion is the proportion of looks on screen. Looks elsewhere were disregarded.

CP and MP trials deviate around 4000 ms into the trial. PTL of CP trials peaks at 4000-5000 ms, and PTL in MP trials has its lowest point around 5000-6000 ms. Resuming to chance is often used as a rationale for setting the end boundary of the post-naming window (Delle Luche et al., 2015). Although PTL is at chance level (0.5) at target onset, it does not resume to chance for the entire trial, so the influence of the target word remains present. Although post-naming windows typically have a duration of 2000 ms, based on PTLs typically returning to chance level within this time frame, longer time windows are not uncommon (for examples, see Delle Luche et al., 2015). A longer time window is better suited in our case, since in our data PTL does not return to chance. We set the end of the post-naming window at 6000 ms. We did not extend the time window to the end of the trial (at 7000 ms), because data towards the end of the trial will be based on fewer children; children often tend to look away from the screen towards the end of the trail and data are therefore less meaningful. We set our post-naming window to 3000-6000 ms.

***Data cleaning***

We employed pre-set exclusion criteria regarding trials, blocks and participants. Criteria are based on Ramachers et al. (2017).

**Test trials** were excluded from final analysis if:

(1) The child did not fixate on both objects during the pre-naming window;
(2) The child looked at the screen for less than 750 ms (25%) of the post-naming window[6];
(3) An equipment error or experimenter error occurred that could not be corrected post-hoc;
(4) Trials that were not completed due to a child not finishing the experiment.

Criterion (1) was administered before plotting the data to determine the post-naming window. Criterion (2) was administered only after it was determined what the post-naming window would be. Criterion (3) was never violated. Two children did not complete the second block of their experiment, which resulted in a few missing trials based on criterion (4). For exact numbers of excluded trials per criterion, see Table 4.4.

**Table 4.4** Overview of excluded trials.

| | Criteria | CP trials | MP trials |
|---|---|---|---|
| **(1)** | **pre-naming window not fixated on both objects** | 92 (8%) | 156 (13.7%) |
| **(2)** | **< 750 ms looking time in post-naming window** | 19 (1.6%)<br>(11 (=1%) also bad pre-naming window) | 32 (2.8%)<br>(20 (=1.7%) also bad pre-naming window) |
| **(3)** | **Equipment or experiment error** | n.a. (0%) | n.a. (0 %) |
| **(4)** | **trials not completed** due to not finishing experiment | 2 (0.2%) | 4 (0.4%) |

A **block** was excluded if:

(1) There was not at least one valid CP trial and one valid MP trial within one block;
(2) Total looking time during learning trials was less than 20 out of the total 60 seconds for both the target learning trials and the distractor learning trials.[7]

In total, 38 blocks (31.3%) were excluded based on criterion (1). No blocks had to be excluded based on criterion (2).

6 Also used in Buckler & Fikkert (2016) and Tsuji et al. (2016).
7 This criterion was taken as a measure of the general attention paid during the learning phase and children who pay more attention to the novel objects during this phase should be better able to retain the novel word-object mapping (Hilton & Westermann, 2016).

A **participant's** data was excluded when:

(1) Subject did not meet the inclusion criteria listed in the Subjects section (5 subjects: 3 had speech therapy, 1 experienced hearing loss; 1 had reduced vision):
(2) Subject refused to participate (2 subjects):
(3) The caregiver interfered with the child's behaviour (1 subject):
(4) The child used a self-conceived label for either one of the novel objects (1 subject):
(5) Both blocks had to be excluded due to criteria mentioned above regarding trial and block exclusion – resulting in no participant data left (5 subjects).

In sum, 14 subjects (6%) were excluded based on these criteria. In the final sample, we could include data of 137 out of the tested 151 participants.



## Analysis

Children's target recognition after the target was named was inferred from the presence of a naming effect, typically measured as an increase in target looking relative to a baseline measure upon hearing the target label (e.g. Delle Luche et al., 2015). There are several ways to compute a naming effect. Similar to Ramachers et al. (2017), we use the PTL measure, which is used in most cases. Within the pre- and post-naming window, PTL is calculated by dividing the total amount of looking time towards the target (T) by the total amount of time a subject spent looking at target or distractor (T+D) within the relevant time window. This is reflected in the following formula:

$$PTL = T/(T+D)$$

The value for PTL lies between 0 and 1. For each trial of each participant we determined the proportion of target looking (PTL) for both the pre-naming (i.e. baseline) and the post-naming window. To calculate a naming effect, the increase in PTL from pre-naming to post-naming window needs to be calculated as follows:

$$\text{Difference score} = [\text{post-naming PTL}] - [\text{pre-naming PTL}]$$

This results in a value in between -1 and 1. When the target word was recognized, this is reflected by a value significantly above zero (i.e. positive naming effect), indicating a larger PTL in the post-naming window than in the pre-naming window. We used difference scores per trial per participant as our dependent variable (henceforth: *naming effect*). Computing naming effects using this post-naming minus pre-naming PTL measure for each individual participant serves to control for possible effects of preference for a particular stimulus (e.g. Ramachers et al., 2017; White & Morgan, 2008; Mani & Plunkett, 2011; Quam & Swingley, 2010). It also controls for a general target preference (Ramachers et al., 2017).

Ramachers and colleagues already confirmed that objects were of similar attractiveness to the children. As our own verification of this, however, we calculated the proportion of looks towards each of the object during the object familiarization phase.

Also, similar to Ramachers and colleagues, we assessed the impact on target preference due to the learning phase, where only the target is named. We calculated PTL in the object familiarization trial (= before learning) as well as the pre-naming window (= after learning) and compared the PTL by means of a paired samples t-test. We controlled for any pre-naming preference by using difference scores between pre- and post-naming as our dependent variable.

A prerequisite for the experiment to be meaningful is that children successfully learned the novel word-object mapping. We calculated naming effects for filler and CP trials and tested whether the naming effects were significantly above zero using one-sample t-tests. We would at least expect a naming effect for filler trials if children are in general performing the task. A CP naming effect would indicate that children successfully learned the novel word-object mapping during the learning phase, and recognize the target word. We tested this for each of the three target vowels.

Next, naming effects were calculated and compared for CP and MP trials. An overall larger naming effect for CP than MP trials indicates that children are in general sensitive to vowel mispronunciations. However, this can differ for different vowel conditions. An MP may show:

(1) **A positive naming effect** (significantly > 0). This means children look significantly more at the target object after naming than before naming. This means that the word is still recognized despite the vowel being mispronounced. If this is the case, we cannot say word recognition is hindered unless the naming effect for the MP is significantly smaller than the naming effect for its CP counterpart (as attested for tone contrasts in Ramachers et al., 2017).

(1) **No naming effect** (naming is not significantly different from zero). This means that children are unsure which object is referred to. Proportion of target looking does not change after naming. This would mean that target recognition is hindered to such an extent that recognition fails (as attested for two- and three-feature segmental MPs in White & Morgan, 2008).

(1) **A negative naming effect** (significantly < 0). This means that children look more at the distractor after naming than before. This is interpreted as that the mispronunciation is noticed and the word is interpreted as a label for the distractor

for which the child has no label yet - the formation of a novel mapping between auditory label and distractor object. The MP is regarded a different word than the CP, so word recognition is hindered. Since we use novel objects, rather than for example White & Morgan (2008) who used familiar objects, this is a plausible outcome (cf. Delle-Luche et al., 2015).

We ran an analysis using mixed effects models in JMP (version 13 – SAS Institute Inc., 1989-2019) which were similar to the models used for the semantic priming experiment in Chapter 3. We used *Naming Effect* (i.e. difference score between post- and pre-naming window) as our dependent variable. As Random Factor, we used *Subject.* As fixed factors we included *Trial Type* (2 levels: CP and MP) and *Direction of Change* (i.e. vowel condition) with six levels (2 directions for 3 vowel pairs). We determined the presence of naming effects for CPs and MPs in all vowel conditions to assess word recognition in the different conditions.



## Results

We assessed attractiveness of the four novel objects. For each object pair, we calculated the proportion of looks based on the object familiarization trial, where the objects are introduced and none of them is named. Children spent 51,4% of looking time towards the yellow object (Object A in Figure 4.2) and 48,6% of looking time towards the purple object (Object B in Figure 4.2). Children spent 45,4% of looking time looking at the pink object (Object C in Figure 4.2) and 54,5% at the blue object (Object D in Figure 4.2). Ergo, no striking differences in attractiveness between the different objects were found.

In familiarization, mean PTL was 0.512, which was not significantly above chance ($t(244) = 1.18$; $SD = 0{,}16$; $p = 0.118$), indicating no target preference before the learning phase. Mean overall pre-naming PTL in test-trials (CP and MP trials) of 0.52 was close to chance, although statistically significantly above chance in a one-sided t-test ($t(820) = 2.36$; $SD = 0.25$; $p = 0.009$), which may indicate the targets to be more attractive than the distractor objects after learning.

Mean PTL values and standard deviations for pre- and post-naming windows per trial type and condition are listed in Table 4.5.

**Table 4.5** Overview of mean proportion of target looking (PTL) in pre- and post-naming windows for correct pronunciation trials (CP) and mispronunciation trials (MP) per vowel condition.

| | | CP | | MP | |
|---|---|---|---|---|---|
| **Vowel Pair** | **Condition** | pre-naming PTL | post-naming PTL | pre-naming PTL | post-naming PTL |
| **/ø/ - /o/** | **/ø/[o]** | 0.62 (0.22) | 0.67 (0.24) | 0.50 (0.28) | 0.49 (0.30) |
| | **/o/[ø]** | 0.53 (0.22) | 0.70 (0.27) | 0.50 (0.28) | 0.55 (0.31) |
| **/e/ - /ø/** | **/e/[ø]** | 0.50 (0.20) | 0.71 (0.25) | 0.53 (0.25) | 0.48 (0.28) |
| | **/ø/[e]** | 0.51 (0.24) | 0.71 (0.25) | 0.49 (0.24) | 0.53 (0.27) |
| **/o/ - /e/** | **/e/[o]** | 0.52 (0.23) | 0.72 (0.23) | 0.50 (0.26) | 0.38 (0.29) |
| | **/o/[e]** | 0.55 (0.25) | 0.68 (0.28) | 0.50 (0.28) | 0.45 (0.29) |

/V/ indicates CP vowel and [V] indicates MP vowel. Standard deviations in parenthesis.

The model revealed a main effect of *Trial Type* (CP vs MP: $F(1,697.3) = 48.8$; $p < 0.0001$) indicating a difference in naming effect for CP versus MP trials, with a larger naming effect for CP trials (0.09 above intercept) (see Figure 4.4). We did not find a significant main effect of *Direction Of Change* (6 levels) ($F(5,366) = 1.22$; $p = 0.148$), but we did reveal a significant interaction of *Trial Type*Direction Of Change* ($F(5,696.5) = 2.23$; $p = 0.049$), implying that the naming effect was affected differently for MPs and CPs for different conditions.

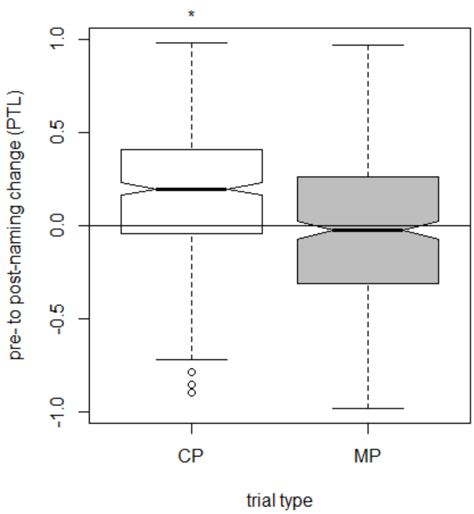



**Figure 4.4** Naming effect, i.e. pre- to post-naming change in proportion of target looking (PTL), for CP and MP trials (overall). CP = correct pronunciation. MP = mispronunciation.

We found a strong naming effect for filler trials (M = 0.24; SD = 0.36; t(824) = 19.2 ; p < 0.0001) indicating that the children were on task. We also found a clear naming effect overall for CP trials (M = 0.16; SD = 0.34; t(429) = 9.8; p < 0.0001), which indicates that overall, children were successful in their novel word-object mapping and during the test phase they recognized the target word they learned during the learning phase. We did not find a significant overall MP naming effect (M = -0.02; SD = 0.41; t(390) = -1,09; p = 0.27 (two-sided)), which indicates MPs generally did not result in target word recognition. Thus, overall, children were sensitive to mispronunciations.

**Table 4.6** Naming effects for CPs of the three target vowels /e/, /o/ and /ø/. ** indicates p < 0.0001.

| Target vowel | CP naming effect estimate (SD) |
|---|---|
| /e/ | 0,21 (0,34)** |
| /ø/ | 0,13 (0,35)** |
| /o/ | 0,14 (0,33)** |



We found significant naming effects for CPs for all three target vowels (see Table 4.6). More detailed analysis of naming effects for MPs for all six vowel conditions showed no significant positive naming effects for any of the conditions. A negative naming effect was found when /e/ was mispronounced as [o] (LSM = -0,11 ; SD = 0.046; p = 0.0072). Results for CPs and MPs of each contrast are summarized in Table 4.7 and presented in the boxplots per contrast in Figures 4.5 (place), 4.6 (labiality) and 4.7 (two-feature contrast).

**Table 4.7** Naming effects; change in PTL from pre-naming to post-naming window.

| Vowel pair | Contrastive feature(s) | Featural Distance | Condition | CP naming effect (SD) | MP naming effect (SD) |
|---|---|---|---|---|---|
| **/ø/ - /o/** | place | 1 | /ø/[o] | 0.13 (0.35) ** | - 0,01 (0.41)<br>p = 0.42 |
| | | | /o/[ø] | 0.14 (0.34) ** | 0.04 (0.43)<br>p = 0.2 |
| **/e/ - /ø/** | labiality | 1 | /e/[ø] | 0.21 (0.34)** | - 0,04 (0.34)<br>p = 0.17 |
| | | | /ø/[e] | 0.13 (0.35) ** | 0.04 (0.37)<br>p = 0.13 |
| **/o/ - /e/** | labiality & place | 2 | /o/[e] | 0.14 (0.33) ** | - 0.05 (0.42)<br>p = 0.15 |
| | | | /e/[o] | 0.21 (0.34) ** | - 0.11 (0.38)*<br>p = 0.0072 |

Positive values indicate target preference. Negative values indicate distractor preference. /V/ indicates CP vowel and [V] indicates MP vowel. CP = correct pronunciation. MP = mispronunciation. For each CP the same CP occurs in two conditions, which are grouped. Thus, the same values occur twice for CP naming effects. * indicates significant difference from zero with p < 0.01. ** indicates significant difference from zero with p < 0.0001. Note that all CPs show significant positive naming effects. MPs do not show (positive) naming effects, apart from when /e/ is mispronounced as [o], which results in a negative naming effect.

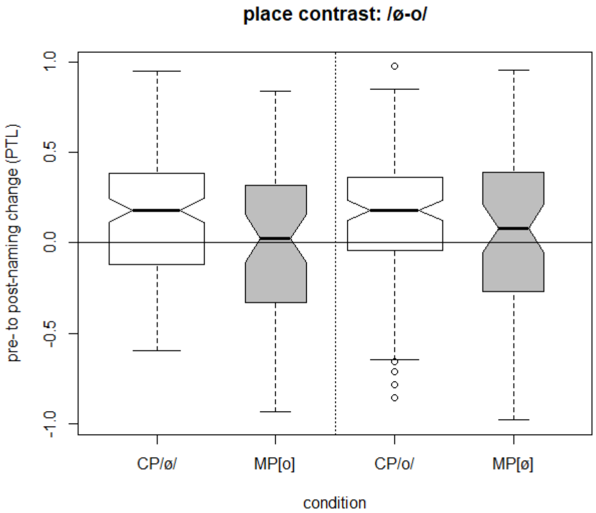



**Figure 4.5** Place contrast. Boxplot showing the median and interquartile range of the pre- to post-naming change (naming effect based on difference score in post-naming and pre-naming PTL). In CP and MP trials for the contrast between /ø/ and /o/. The left panel shows results for a coronal CP and its dorsal MP. The right panel shows results for a dorsal CP and its coronal MP. White boxes indicate CP trials and grey boxes indicate MP trials. Width of the box is proportionate to the sample size. Whiskers show the range between lower and upper quartile. Dots represent outliers. The solid horizontal line indicates chance level.

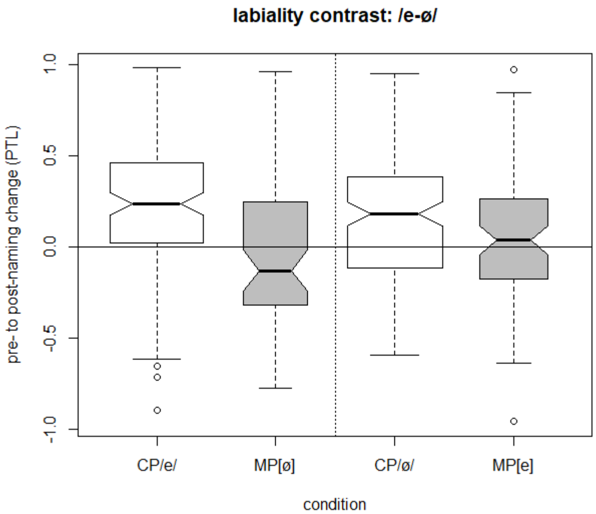



**Figure 4.6** Labiality contrast. Boxplot showing the median and interquartile range of the pre- to post-naming change (naming effect based on difference score in post-naming and pre-naming PTL). In CP and MP trials for the contrast between /e/ and /ø/. The left panel shows results for a nonlabial CP and its labial MP. The right panel shows results for a labial CP and its nonlabial MP. White boxes indicate CP trials and grey boxes

indicate MP trials. Width of the box is proportionate to the sample size. Whiskers show the range between lower and upper quartile. Dots represent outliers. The solid horizontal line indicates chance level.

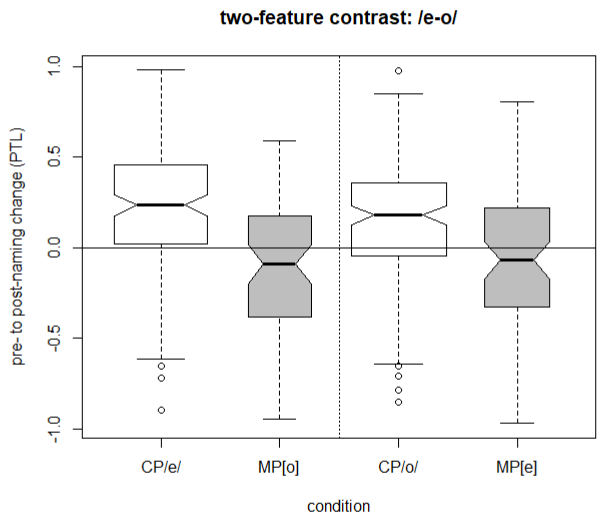



**Figure 4.7** Two-feature contrast. Boxplot showing the median and interquartile range of the pre- to post-naming change (naming effect based on difference score in post-naming and pre-naming PTL). In CP and MP trials for the contrast between /e/ and /o/. The left panel shows results for a coronal nonlabial CP and its dorsal labial MP. The right panel shows results for a dorsal labial CP and its coronal nonlabial MP. White boxes indicate CP trials and grey boxes indicate MP trials. Width of the box is proportionate to the sample size. Whiskers show the range between lower and upper quartile. Dots represent outliers. The solid horizontal line indicates chance level.

Results from two additional analyses – comparing naming effects for CP and MP conditions, and comparing post-naming PTL for CP and MP conditions – are included in Appendix 7 and 8.

## Discussion

To complement our studies in Dutch adults, we investigated the perception of place and labiality vowel contrasts in 38-month-old children acquiring Dutch.

We used a word learning variant of IPLP, adopted from Ramachers and colleagues (2017). We tested the toddlers' perception of a three-way vowel contrast between coronal, coronal labial, and dorsal labial vowels at a lexical level, using the contrast between tense vowels /e, ø, o/. Thus, we investigated the same three contrasts as we looked at in previous chapters in adults:

(1) One-feature contrast (place): coronal versus dorsal - /ø/ versus /o/;
(2) One-feature contrast (labiality): labial versus nonlabial - /e/ versus /ø/;

(3) Two-feature contrast (place & labiality): /e/ versus /o/.

Assuming three-year-old children would have adult-like phonological representations of the features of interest, we predicted that this perception study would show similar directional asymmetries as found in adults in Chapter 2. More concretely, we predicted asymmetries for the single-feature contrasts: a place asymmetry and a labiality asymmetry. For the two-feature contrast, both directions of change were predicted to result in phonological conflict.

Results demonstrate that the toddlers successfully learned the novel words regardless of which vowel was used – as indicated by positive naming effects for CPs for all three target vowels. Overall, naming effects were smaller for MPs than for CPs, indicating that children were sensitive to word-medial vowel mispronunciations. This MP-effect also supports successful build of representations of CPs.

Results show no significant positive naming effects for any of the MPs, indicating that regardless of which phonological feature changed, children did not recognize the target word when the vowel was mispronounced. Word recognition was hindered by the MP in all vowel conditions. This demonstrates that three-year-old Dutch children are sensitive to both place and labiality mispronunciations in vowels.

Below, we will discuss the results and their implications in more detail. We will discuss to what extent the results are in line with our expectations, but also compare current findings to predictions based on other models, i.e. the predicted directional asymmetries based on NLM (Kuhl, 1991; Kuhl et al., 2008), and the NRV framework (e.g. Polka & Bohn, 2011; Masapollo, Polka, & Molnar, 2017; Masapollo, Polka, & Ménard, 2017). A more elaborate discussion comparing children and adult vowel perception is part of Chapter 5 – General Discussion.

**Lexical Encoding of Place and Labiality Features of Vowels in Dutch Toddlers**

The finding that children were successful at learning all novel word-object mappings in the current experiment implies that unlike Dutch 14-month-olds (production: Fikkert & Levelt, 2008; perception: Fikkert, 2010) 38-month-olds no longer assume [CORONAL] and [LABIAL] to be mutually exclusive – at least in vowels. In particular, the finding that CPs of words with coronal labial vowel /ø/ are successfully learned and recognized supports this notion. It implies that when the word is heard, no SR features are conflicting with its UR features, so the coronal SR feature of the input with [ø] does not contrast with its [LABIAL] UR feature. This is in line with what we would expect if three-year-olds have adult-like representations.

We predicted directional asymmetries in perception of mispronunciations of labiality and place. Predictions regarding naming effects for MPs and directional asymmetries are recapitulated in Table 4.8, along with the experiment outcomes.

**Table 4.8** Predicted naming effects for MP for each condition.

| | | | | Predictions | | | Results | |
|---|---|---|---|---|---|---|---|---|
| **Vowel pair** | **Contrast** | **CP** | **MP** | **Mapping** | **Naming effect** | **Asymmetry** | **Naming effect** | **Asymmetry** |
| **/ø ~ o/** | place | /ø/ | [o] | no-mismatch | yes | yes | no | no |
| | | /o/ | [ø] | mismatch | no/negative | | no | |
| **/e ~ ø/** | labiality | /e/ | [ø] | *aborted* | no/negative | yes | no | no |
| | | /ø/ | [e] | no-mismatch | yes | | no | |
| **/o ~ e/** | place & | /o/ | [e] | mismatch | no/negative | no | no | ? |
| | labiality | /e/ | [o] | no-mismatch | no/negative | | negative | |
| | | | | *aborted* | | | | |

CP = correct pronunciation. MP = mispronunciation. Surface representation of MP is assumed to map onto the underlying representation of the CP.



For both single-feature contrasts, we predicted asymmetrical patterns of sensitivity to MPs, but we did not reveal such asymmetries. For the single-feature contrasts, all MPs show no naming effect, indicating no recognition of the canonical word. This implies that all MPs hindered word recognition, which was against our expectations. We predicted that in both single-feature contrasts one of the MPs would not hinder lexical access. The findings indicate that 38-month-old Dutch toddlers are sensitive to place and labiality contrasts. This is in line with our expectation that children have acquired the three-way distinction – similar to adults. However, absence of asymmetries is against our expectations, and different from the adult findings.

For the two-feature contrast, we found no naming effect when /o/ was mispronounced as [e], and we found a significant negative naming effect when /e/ was mispronounced as [o]. This implies that as we predicted, word recognition was hindered in both directions. The negative naming effect indicates that rather than uncertainty about which is the target object (i.e. when naming effect is not different from zero), children interpret the MP[o] as a label associated with the distractor object. The MP was interpreted, beyond doubt, to be a different word than the CP/e/ and hence referred to the distractor object, for which they did not know a label yet. We assumed this negative naming to be a reasonable outcome, since adults showed an asymmetrical pattern in mismatch negativity (MMN) response; a change from /e/ to [o] was perceptually larger than vice versa – albeit nonsignificant. Thus, children show a similar pattern, where both MPs seem to result in phonological conflict, but a change from /e/ to [o] may be an even stronger conflict than vice versa. The difference between the naming effects for both MPs in the two-feature contrast, however, was not significant, and we can hence not conclude that perception is asymmetrical.

**Alternative theories**

Also, based on a frequency account like the Native Language Magnet Theory (NLM), and saliency account like the Natural Referent Vowel (NRV) framework, asymmetries would be expected for the single-feature contrasts. Both NRV and NLM predict perceptual asymmetries based on phonetic discrimination rather than on phonological processing. In Table 4.9 predictions from these models and our own hypothesis are stated, and contrasted with current outcomes. NLM in particular would predict asymmetries for single-feature contrast opposite to what we predicted. Yet, we found no asymmetries in either of the single-feature contrasts. Current data does not fit any of these models. Reasons for why we do not find asymmetries for these contrasts will be revisited below in in the next section. We will now turn to the third contrast we tested: the two-feature contrast between /e/ and /o/.

Based on frequency distributions, NLM predicts an asymmetry for the two-feature contrast: a mispronunciation of /o/ as [e] would be easiest to discriminate, because [e] is a more frequent sound in Dutch than /o/. This is opposite to our adult findings, and not in line with child data either, as children show no asymmetry. Moreover, although it was a nonsignificant pattern, the MP of /o/ as [e] appeared less salient than vice versa, which is in opposite direction from NLM's predictions.

**Table 4.9** Predicted directional asymmetries in change detection of labiality and place features of vowels made by different models on infant vowel perception for one-feature contrasts of place (/ø-o/) or labiality (/e-ø/) and the two-feature contrast of place and labiality (/e-o/).

| model | | place | labiality | place & labiality |
|---|---|---|---|---|
| **NLM** | | /ø/ → /o/ | /e/ ← /ø/ | /e/ ← /o/ |
| **NRV** | F1-F2 | /ø/ → /o/ | /e/ → /ø/ | /e/ → /o/ |
| | F2-F3 | /ø/ ← /o/ | /e/ ← /ø/ | /e/ ← /o/ |
| **Current hypothesis** | | /ø/ ← /o/ | /e/ → /ø/ | /e/ - /o/ |
| **Current results** | | /ø/ - /o/ | /e/ - /ø/ | /e/ - /o/ |

The arrow indicates the direction of change that is easiest to detect. Results of the current experiment are included.

The NRV framework does not make clear predictions regarding an asymmetry for the two-feature contrast, since both /e/ and /o/ are similar in how peripheral they are. /o/ is more focalized than /e/ in F1-F2, but less focalized than /e/ in F2-F3. Whether or not to expect an asymmetry thus differs depending on F1-F2 or F2-F3. Symmetrical results neither confirm nor reject this hypothesis.

In the two-feature contrast (/e-o/) proportion of target looking over time shows a larger PTL drop after naming for MPs in the two-feature contrast (/e-o/) than either of the single-feature contrasts - see time-course graphs in Appendix 9 (not statistically tested). Previous literature suggests that toddlers can show graded sensitivity to MPs (White and Morgan, 2008), with larger sensitivity when MPs that differ regarding two features compared to MPs that differ from CP regarding a single phonological feature. However, we hypothesized only one feature to be conflicting in either direction of change for this contrast. For this reason, it is not evident that the MPs are regarded as two-feature contrasts in perception. Alternatively, it is possible that acoustic distance plays a role, which is larger in this contrast than in the single-feature contrasts. This could constitute a more salient difference between the sounds as well. If so, the phonetic/acoustic level of processing still impacts perceptual difference between CP and MP, and may be stronger than phonological processing in some cases.



***Does no asymmetry mean symmetry?***

In sum, our hypothesized underspecification account, NRV, and NLM all predicted perceptual asymmetries for the single-feature contrasts. We do not find evidence of perceptual asymmetries in the current experiment.

An alternative model to consider is the TRACE model, a computational model of speech perception (McClelland & Elman, 1986). Alternative accounts of spoken word recognition have stressed the role of phonological neighbours (e.g. Culter, 1995) or the role of cohort competitors (e.g. Marlsen-Wilson & Welsh. 1978). The TRACE model includes a role for both these aspects in a continuous mapping model. In TRACE, a listener evaluates speech input against a set of lexical candidates for recognition. If applied to infant or child word recognition, the infant's lexicon thus impacts the sensitivity to words. Both cohort competitors and phonological neighbours play a role. Words in the lexicon compete at the level of their phonological overlap.

Mayor and Plunkett (2009) used the TRACE model to study infant sensitivity to vowel and consonant mispronunciations. They simulated IPLP experiments using jTRACE (Strauss, Harris, & Magnuson, 2007), a reimplementation of the TRACE model, using typical lexicons for British 15-, 18-, and 24-month-olds. They tested sensitivity to word-initial consonant mispronunciations and word-medial vowel mispronunciations. TRACE predicts that sensitivity to lexical mispronunciations will be influenced by competition between lexical items. Results from Mayor and Plunkett (2009) imply that the sensitivity to onset consonant mispronunciations is directly influenced by the number of cohort competitors, a by-product of the increasing size of the lexicon. The sensitivity to medial vowel mispronunciations is, in contrast, a subtle effect of co-activation of neighbourhood words enhanced by the forced choice procedure. This seems unrelated to the mispronunciation being a vowel per se, but rather the mispronunciation not being the word initial sound.

In computational models like the TRACE model, lexical entries from dense neighbourhood sizes should exhibit greater mispronunciation effects than those with sparse neighbourhoods (see also Mayor and Plunkett, 2014). Based on the Lexilijst Nederlands (Schlichting & Lutje Spelberg, 2002) current findings without asymmetries for word-medial vowels in toddlers does not contradict these predictions based on the TRACE model.

Another logic relates frequency to a transition from acoustically detailed and potentially exemplar-based representations to more robust lexical representations (see Mani and Plunkett, 2010; Mayor and Plunkett, 2014). So, in addition to vocabulary size, word familiarity may affect infant spoken word recognition. Word frequencies would predict symmetrical results, since we used novel words and hence familiarity is similar for all MPs.

In the General Introduction (Chapter 1) we also discussed some adult-based models that do not predict asymmetries for native between-category perception of vowel changes. Some views assume that all phonological features are specified in our representations (full specification; e.g. Chomsky & Halle, 1968). Others assume that we store most if not all phonetic detail as a cloud of exemplars, which form our lexical representations (e.g. exemplar based models; e.g. Goldinger, 1998; Johnson, 1997; Ernestus, 2014) – similar to Mani and Plunkett (2010). Although such models might explain perceptual asymmetries in terms of, for example, differences in frequencies of occurrence, both full specification and exemplar-based views would in principle predict symmetrical native between-category perception of phonological contrasts. Thus, these models seem to agree with current results.

Nonetheless, several studies indicate that perceptual data of younger Dutch children supports coronal underspecification (e.g. consonants and vowels: Fikkert, 2010; consonants: Tsuji et al., 2016). Coronal underspecification is also supported in Dutch adults (Chapter 2). It would be incongruous to assume full or exemplar-based specification for 38-month-olds, but not for younger toddlers nor for adults.

Alternatively, 38-month-olds may in fact be so adult-like that the task is too easy for them, and consequently the paradigm is not sensitive enough to detect perceptual asymmetries. As also noted by Tsuji et al. (2016), adult listeners would certainly be able to detect a change in a simple and straightforward mispronunciation task such as ours, which is why adult studies predominantly assessed responses in more sensitive paradigms such as cross-modal (semantic) priming or EEG (e.g. Eulitz & Lahiri, 2004; Chapter 2 and 3; Scharinger et al., 2012; Scharinger & Lahiri, 2010; Kotzor et al., 2015). Semantic priming / lexical decision tasks are cognitively more demanding, and EEG experiments tap into processing directly, since no behavioural response is required. Indeed, Van der Feest (2007) reports adult sensitivity to mispronunciations in IPLP regardless of direction of change. If children have acquired the place and labiality contrasts at an adult-like level, and task demands are low for them, this may be the reason why we found sensitivity to all directions of change rather than directional asymmetries.

Using IPLP, Altvater-Mackensen and colleagues (2014) found asymmetrical perception of manner of articulation (consonants) in Dutch 18-month-olds, but not in 24-month-olds, who were sensitive to both directions of change. This does not mean that 24-months-olds and older children do not show perceptual asymmetries in IPLP in general, since Van der Feest and Fikkert (2015) only found voicing asymmetries in Dutch 24-month olds and not in 20-month olds, who did not show voicing sensitivity in any direction yet. Altvater-Mackensen and colleagues' findings suggest that sensitivity to different contrasts emerges at different times over the course of development. That 24-month-olds are sensitive to both directions of change in a manner contrast, whereas they demonstrated asymmetries at a younger age, supports the possibility that toddlers in the current study are already highly proficient in discriminating the contrast and no longer show asymmetrical patterns in IPLP which may be present at a younger age.



In sum, it is difficult to draw conclusions regarding phonological representations used by our target group based on current findings. Indeed, if the task is not sensitive enough for this age and contrast, it is possible that perception of the toddlers is asymmetrical. This could be similar to adults. The fact that three-year-olds distinguish the three-way contrast at a lexical level does not prove that toddlers do so the same way as adults do.

**Limitations and methodological considerations**

IPLP (with or without word learning, using either familiar or novel objects) has not often been used in our age group, let alone for testing vowel quality contrasts. In fact, sensitivity to vowel features has not been tested at a lexical level in the target age group before. From the current study we learned that perhaps this paradigm was too easy for the current contrast in the current age group. The fact that the task was not too easy for an overlapping age group in the study of Ramachers et al. (2017) tells us that the higher proficiency in our study is probably related to the phonological skills of the participants. Hence, the study shows that 38-month-olds acquiring Dutch are very proficient and sensitive to the three-way contrast between coronal, coronal labial, and dorsal labial vowels. Our hypothesis that children at this age are adult-like with respect to these phonological features is not rejected so far.

Apart from using more sensitive methods like EEG, a more sensitive analysis of IPLP looking behaviour could have been an alternative. We considered doing a reaction time (RT) analysis in addition to the current time-window based total looking time analysis. However, for an RT analysis, only trials where a child is not looking at the target at the time point when the target is named can be included, so a gaze shift towards the target can be measured as RT. Since children looked roughly 50/50 towards the target/distractor during the pre-naming window, we would lose about half of the currently included trials. Furthermore, detailed information about the fixation of the child at time of naming is necessary to determine which trials can be included. A reliable measures of eye gaze at each time point is required to do this well. Unfortunately, video recordings of half of the participants were corrupted with respect to frame timing, which would render it impossible do a reliable offline frame-by-frame coding of looking behaviour. Our online coding is not precise enough for an RT analysis, because exact time points are important for determining the child's RT (rather than also the RTs of the experimenter) and the child's gaze direction during naming. This was not a problem for the current time-window analysis, because longer time-windows were used. In sum, it was not feasible to run a reliable RT analysis using the current data set.

Furthermore, it could have been useful to have more information about the stage of acquisition of the individual participants with respect to whether or not they indeed produce coronal labial vowels. There was a lot of variance in the data. Perhaps not all children were at the same stage regarding their use of coronal labial vowels in production. Knowledge about the individual children's stage of vowel development in production would have enabled us to split the data for different developmental groups and see how this affected the patterns in their perception. N-CDI-data perhaps would have aided our understanding of the data.

Another factor that perhaps increased the amount of variance in our data is the use of a within-subject design: each child was tested on both directions of a single contrast. Investigating directional asymmetries, a within-subject design is suitable, since the same individual then provides data for both directions and general individual differences are cancelled out. However, using a within-subject design increased duration of the experiment compared to only testing a single direction per child. Blocks that did not provide enough good trials were most often the second test block. It was not feasible to only include data from participants that had enough trials in both their blocks without losing too much data. Although a within-subject design is elegant when testing directional asymmetries, and limits the required number of participants, unfortunately we did not end up with a pure within-subject data set.

It may be that children's attention spans were not long enough to fully engage in the experiment for its entire duration. Potentially low task demands may have resulted in boredom effects in the second block. Children in Ramachers et al. (2017) had more difficulty perceiving the contrast of interest, which may have increased task demand. This may explain why Ramachers and colleagues were able to only include data of children that had a sufficient number of good trials in both blocks.

We noticed that naming effects for both CPs and MPs were smaller (i.e. closer to zero) in the second block compared to the first block. Stronger effects would have been found in only the first block. Since we controlled for order effects by using several different experiment versions, and counterbalanced order of directions, the difference between block 1 and block 2 did not impact our results with respect to the effects for particular conditions.

Another methodological consideration is the use of familiar or novel words. The fact that we found no naming effects for any of the MPs is probably related to the fact that we used novel words and objects. In a paradigm where familiar objects/words are used, a mispronunciation is still most similar to the word for the target object and less similar to the word for the distractor object. Hence, mispronunciations are not considered to refer to the distractor. As a consequence, studies using familiar objects often find naming effects for mispronunciations (e.g. Van der Feest & Fikkert, 2015). In such cases, a comparison between CP and MP is called for (= MP-effect). Since we did not find naming effects for mispronunciations, we could already conclude that they did not lead to word recognition. A comparison of CP and MP to assess whether MPs hindered word recognition was not necessary.



**Further research**

Interestingly, we found a larger and more stable naming effect for CP words with /e/ as the target vowel compared to both /ø/ and /o/, which are labial. It would be interesting to run more experiments to assess how phonological features impact learnability. Are labial vowels more difficult to learn / build a representation for than nonlabial vowels? Or does learnability have more to do with the combination of features of the vowel and the consonants in the word? It has been argued that children's early production patterns favour particular CV patterns (Davis & MacNeilage, 1995). Preferred CV co-occurrences have been reported in children's early production in many different languages (Davis, MacNeilage & Matyear, 2002; Kern & Davis, 2009; Kern, Davis & Zink, 2010). In these preferred CV co-occurrences, vowels and consonants have the same place feature, so for example coronal consonants co-occur with coronal vowels. This co-occurrence has been reported for the majority of languages, and this is also found in Dutch production data reported by Fikkert and Levelt (2008). In both our word templates /tVs/ and /dVn/ all consonants were coronal. Perhaps words with /e/ were easier to learn because all phonemes in the word were coronal, whereas for the vowels /ø/ and /o/, words had labial and/or dorsal features of the vowel combined with coronal consonants. Perhaps the latter two word forms are therefore relatively complex and more difficult to learn than the former. For the vowels of interest, future studies could look into the learnability of novel words in perception depending on the CV feature structure, similar to, for example, Van der Feest, Fikkert, and Davis (2016), who investigated CV interactions in word recognition, looking at mispronunciations of word-initial consonants (coronal vs. dorsal). They showed that word-initial CV combinations influence word recognition when the place of articulation of only the word-initial consonant is mispronounced. For Dutch-learning children, mispronunciations of word-initial place of articulation hindered word recognition only before back vowels, but not before front vowels.

Further research on the current topic could employ more sensitive methods, tapping into processing directly by using EEG, for example. This would provide greater insight into whether or not perception of place and labiality contrast in vowels is indeed symmetrical in 38-month-old Dutch toddlers. At this age it is important to tap into their lexicon since discrimination and word recognition may not both tap into phonology, an N400 experiment would be preferred over an MMN experiment, although they could also complement each other: do they both reveal different outcomes at this age or not?

In addition, it would be interesting to track development from infant to toddler, to see when exactly children learn that [CORONAL] and [LABIAL] are not mutually exclusive and to gain more insight into how things change when coronal labial vowels enter the system. One could investigate whether there is a particular age window at which children show asymmetries for the current vowel contrasts of interest using IPLP, or other more suitable methods for the youngest infants. It would be interesting to track development from phonetic to phonological processing, investigating the impact of the developing lexicon on the (re)organization of place and labiality in vowels.

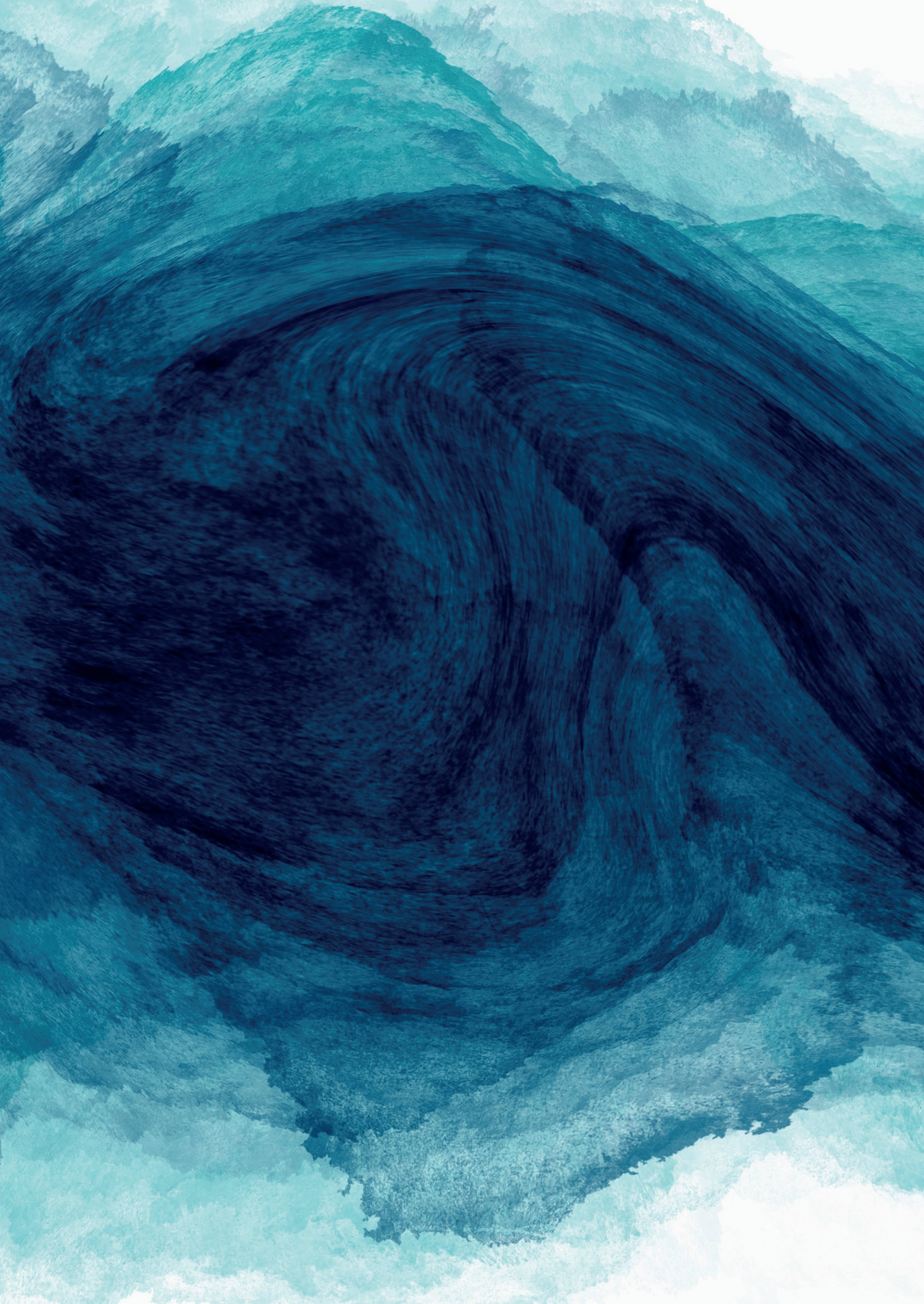
Methods that tap into an individual's perception in a reliable way would have been useful for linking background information about the children to individual differences in task performance. In the current experiment, at most two MP trials per tested contrast were available per child. This does not provide sufficient data to provide insight at the level of individual performance. We were restricted to look at group data. Even though EEG measures may be more sensitive at tapping into processing more directly rather than overt behaviour, EEG is not necessarily reliable at the individual level, even in adults. Also, an EEG experiment like we performed in adults in Chapter 2 would not be feasible in toddlers, because the experiment takes roughly two hours of testing per participant. This is not feasible with toddlers as they have to sit still to avoid muscle artifacts. In general, the field of language acquisition research would benefit greatly from methods that are more sensitive and reliable at the individual level. Unfortunately, such methods are not currently available, but it would be of great value to develop such methods in the future.

**Conclusion**

As the 38-month-old Dutch toddlers in the current experiment demonstrated strong sensitivity to all featural changes (MPs) regarding place and labiality in vowels, we conclude that they are sensitive to both the language general place contrast and the language specific labiality contrast in vowels at a lexical level. They have integrated coronal labial vowels in their phonology, and can keep them apart from both coronal (nonlabial) and dorsal labial vowels. Data imply that the children have incorporated in their phonology that the features [CORONAL] and [LABIAL] are not mutually exclusive. They have acquired a three-way contrast for Dutch tense mid vowels.

Despite these children's high proficiency, we cannot draw strong conclusions regarding the adult-likeness of the phonological representations they employ. Since the current task is considered not sensitive enough to find asymmetries in adult participants either, the lack of asymmetries in toddlers cannot be taken as proof of symmetry without further testing. Thus, more research is required to establish this.

To see how the phonological system of a child acquiring Dutch changes when coronal labial vowels enter their phonology, several approaches might be helpful in the future. We could investigate vowel perception in the same age group using more sensitive methods to get a more detailed view on perception in this age group. This could be complemented by including younger age groups. This might reveal from what age onwards children are able to discriminate all these contrasts at a phonological level, and how representations change throughout development from infant to toddler.



# Chapter 5

## General discussion and conclusion

The central topic in the current dissertation is how a language user's mind deals with phonological contrasts, a prerequisite to speech recognition. The conducted experiments were set up to investigate how the language universal contrast between coronal (i.e. front) and dorsal (i.e. back) vowels, as well as the language specific contrast between labial (i.e. round) and nonlabial (i.e. non-round) coronal vowels, are represented in a language where both are lexically contrastive. I aimed to answer the following three research questions:

RQ1. How is the vowel contrast between CORONAL and DORSAL (i.e. place) represented?
RQ2. How is the vowel contrast between LABIAL and nonlabial (i.e. labiality) represented?
RQ3. How do place and labiality interact when both are contrastive in vowels?



Language users' perception of speech was used to study the underlying phonological representation. I collected perception data of Dutch vowel contrasts, using three different methods:

(1) Mismatch Negativity (MMN), measuring electrical brain responses through electroencephalography (EEG);
(2) Semantic priming, measuring reaction times;
(3) A word learning variant of the intermodal preferential looking paradigm (IPLP), measuring looking behaviour.

All three methods have successfully been used in the past to investigate phonological representations (e.g. Eulitz & Lahiri, 2004; De Jonge & Boersma, 2015; Scharinger et al., 2010; Scharinger et al., 2012; Roberts et al., 2014; Kotzor et al., 2015; Scharinger & Idsardi, 2014; Van der Feest & Fikkert, 2015; Tsuji et al., 2016; Ramachers et al., 2017).

## Summary of chapters

### Chapter 2: Adult MMN study

In Chapter 2, I conducted an MMN experiment with Dutch adult listeners. I investigated perception of place and labiality contrast in tense mid Dutch vowels /e/, /o/ and /ø/ using a passive oddball paradigm. MMN is understood to reflect perceptual difference between sounds (e.g. Näätänen et al., 1997). Large MMN responses may indicate phonological conflict.

As predicted, the experiment showed an asymmetry with respect to the contrast between dorsal /o/ and coronal /ø/; the MMN was enhanced when the deviant was coronal. This asymmetry supports the hypothesis that CORONAL is underspecified and DORSAL is specified. In addition, there was a labiality asymmetry (/e/ vs. /ø/); the MMN was enhanced when the deviant was labial. This was not predicted.

The two-feature contrast (/e/ vs. /o/) showed a larger MMN when the deviant was /o/. However, this asymmetry did not reach significance. This is not in line with the prediction that a coronal deviant – in this case /e/ - would lead to an enhanced MMN when the standard is dorsal. The MMN was large in either direction. A large MMN may imply a conflict. A phonological conflict when the deviant is labial (/o/) is in line with the results for the labiality contrast between /e/ and /ø/, where the labial deviant elicited an enhanced MMN. An explanation for the lack of a significant asymmetry may be that place and labiality contrasts cancel each other out: both yield a phonological conflict, but in opposite directions of change. Another explanation is that the two vowels are acoustically very distinct and easy to discriminate.

The experiment replicated German findings regarding place, but yielded different results regarding labiality. I argue that the findings are in line with the hypothesized geometry, and furthermore indicate an important role for the nodes in the feature geometry in vowel processing. This will be discussed in more detail in the section Explaining Asymmetries below.

**Chapter 3: Semantic priming**

In Chapter 3, a cross-modal semantic priming study was conducted to assess whether the unexpected labial asymmetry found in the MMN study in Chapter 2 would be replicated using an entirely different experimental method. In the semantic priming experiment, the same vowel contrasts between tense vowels /e, o, ø/ were tested in words. To investigate whether findings from tense mid vowels generalize to different vowels constituting similar phonological contrasts lax counterparts /ɛ, ɔ, ʏ/ were also included. Given that lax vowels in Dutch are not diphthongized, comparing perception of contrasts in tense and lax vowels could illuminate whether or not diphthongization plays a role in why Dutch listeners show different perceptual patterns than German listeners.

Overall, there was significant priming for correctly pronounced words (CP priming). However, after data cleaning only a limited data sample was left (see Chapter 3 for details). This may explain why not all individual vowel conditions reached significance regarding CP priming. The limited dataset and the lack of CP priming in some conditions rendered it impossible to interpret mispronunciation (MP) priming in a meaningful way. There was no significant MP priming in any condition, presumably due to the small final dataset. Hence, patterns rather than statistical effects were assessed. The patterns were in line with a place asymmetry similar to what we found in the MMN study; there was less priming when the dorsal was mispronounced as coronal than vice versa. This was found both in the tense and lax vowels. Hence, patterns in the data are in line with coronal underspecification and dorsal specification.

Unfortunately, the results remained inconclusive regarding both the labiality contrast and the two-feature contrast. The comparison of tense versus lax vowels remained inconclusive as well.

**Chapter 4: Intermodal Preferential Looking Paradigm in toddlers**

Apart from studying adult perception (Chapter 2 and 3), understanding how the system is built during acquisition may provide insight into the adult system as well. The literature on production implies that the first contrast typically acquired, is the contrast between labial and coronal consonants (e.g. Jakobson, 1941/68; Levelt, 1994; Fikkert and Levelt, 2008). The first vowels that enter the system are the cardinal vowels /a/ - /i/ - /u/ back (e.g. Beers, 1995), indicating a distinction between high and low and between front and back. At this early stage labial and coronal are still mutually exclusive. In the adult state, labial and coronal should not be considered mutually exclusive anymore when it comes to Dutch vowels; labial and coronal may combine in vowels like /ø/. Typically, such front round vowels are acquired rather late. In Chapter 4, the Intermodal Preferential Looking Paradigm (IPLP) was used to test vowel perception in Dutch toddlers at the earliest age when front rounded vowels are typically produced, i.e. around 3 years of age (based on Beers, 1995). Presumably, at this age, children should be able to perceptually distinguish coronal labial vowels (like /y/ and /ø/) from coronal (/i/ or /e/), and dorsal labial vowels (/u/ or /o/) at the lexical level. It is to be expected that the place contrast (coronal vs. dorsal) and the labiality contrast (labial versus nonlabial) in vowels may be present in the child's phonology around this age to account for the three-way contrast. However, coronal labial vowels may not yet be incorporated in phonology in an adult-like fashion. Similar outcomes to adult asymmetries in Chapter 2 are not evident.

5

In this experiment, the contrasts between tense mid vowels /e, ø, o/ were tested. Children learned several novel words (i.e. word-object mappings) with tense mid vowels /e, ø, o/ and were tested on their sensitivity to mispronunciations by measuring the proportion of target looking time when hearing correctly or mispronounced versions of the words they had learned.

The experiment showed a significant naming effect for correctly pronounced words (CPs), indicating that the children successfully learned and recognized the words and participated in the experiment as expected. When the vowel was mispronounced (MPs), no naming effects were found, indicating no word recognition. Absence of a positive naming effect for mispronunciations in all vowel conditions indicates children were sensitive to all mispronunciations of both place and labiality contrasts.

For the two-feature contrast between /e/ and /o/, a negative naming effect was found when a nonlabial vowel was mispronounced as a labial vowel, but not vice versa. This negative naming effect indicates that the MP was considered a different word than the CP. Children responded as if it was the label for the distractor item, and looked more at the distractor than the target object. When a labial was mispronounced as a nonlabial, however, no negative naming effect was found. Although this difference between direction was not a significant asymmetry, the direction of this difference is in line with the direction of the labiality asymmetry revealed for adults using MMN in Chapter 2.

Successful word learning in the IPLP experiment showed that, in perception, three-year-old Dutch toddlers do not consider LABIAL and CORONAL to be mutually exclusive, which supports the hypothesized geometry. They are highly sensitive to all featural changes tested. This may indicate adult-like representations. However, different to our findings for adults in Chapter 2 (MMN), no asymmetries were revealed. Potential reasons for this were extensively discussed in Chapter 4. I will briefly revisit this in the section Comparing children and adults (p. 166)

## Explaining asymmetries

In Chapter 1, I argued for the feature geometry depicted in Figure 5.1 below. This geometry separates coronal and dorsal features from labial features. The geometry reflects mutual exclusivity: [CORONAL] and [DORSAL] are mutually exclusive, whereas either of them may be combined with [LABIAL]. I adopted the underspecification account of the Featurally Underspecified Lexicon (FUL) model (e.g. Lahiri & Reetz, 2002; 2010; Lahiri, 2018) in order to make predictions for perception using the geometry in Figure 5.1.

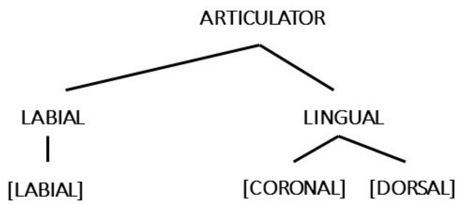



**Figure 5.1** Hypothesized geometry for place of articulation features.

In all three experiments I tested perception of the same Dutch vowel contrasts and assessed whether they yielded asymmetries, i.e. whether within a single contrast one direction of change yielded a larger perceptual difference than the reverse direction of change. When phonological features are conflicting, this is assumed to yield a larger perceptual difference than when there is no phonological conflict. Assuming underspecification of [CORONAL] and specification for [LABIAL] and [DORSAL], but no specification for nonlabial (and no feature that is extractable from the signal), I predicted the following for the three central research questions:

**RQ1 – place contrast:**
[CORONAL] surface representation mismatches with [DORSAL] underlying representation, but not vice versa, predicting a perceptual asymmetry.

This is identical to what would be predicted in the traditional FUL model which does not separate labial and lingual features in the geometry. The prediction is compatible with current results. A clear place asymmetry was revealed by the EEG experiment described in Chapter 2. A similar pattern was found in the semantic priming experiment in Chapter 3.

**RQ2 – labiality contrast:**
A labial surface feature cannot mismatch with a nonlabial underlying representation. A nonlabial surface form, on the other hand, will not trigger any mapping of labiality features. As a result, no mismatch would occur in either direction of change, predicting no asymmetry.

Yet, as already discussed in Chapter 1, this is problematic, as it is not clear how these contrasts could be kept apart. Moreover, it is not compatible with my data, which particularly in Chapter 2 show a clear asymmetrical pattern.



**RQ3 – two-feature contrast:**
As CORONAL is underspecified and DORSAL is specified, the contrast between [CORONAL] nonlabial vowels and [DORSAL, LABIAL] vowels is predicted to result in an asymmetry in the same direction as for a contrast between [CORONAL] and [DORSAL] vowels when both are LABIAL (RQ1). Labiality does not have an asymmetric impact – similar to predictions for RQ2, but the same issue as mentioned above is at stake in this situation as well.

For labiality, I noted earlier that neither direction of change would lead to phonological conflict, which is puzzling in a language where the contrast between labial and nonlabial vowels plays such an important role. Hence, I questioned whether this prediction would indeed be supported by data in perception experiments with Dutch listeners.

A difference between LINGUAL and LABIAL is that vowels are always lingual but not always labial. Vowels are always either [CORONAL] or [DORSAL] so their parental node LINGUAL is always present, even when coronal is underspecified. In contrast, nonlabial vowels not merely lack the feature [LABIAL], but – following the logic employed in the FUL model – the LABIAL node would be absent for these vowels as well. In Chapter 1, I suggested that perhaps not only the features, but also the nodes may play a role in distinguishing between labial and nonlabial vowels. The difference between the two sets of contrasts is that the contrast between coronal and dorsal is language universal (i.e. it is present in all languages), but the difference between labial and nonlabial is language specific, and in many languages, labiality is not contrastive. Rather, back (dorsal) vowels are by default round, and front (coronal) vowels are by default unround.

As mentioned above, adult MMN data as well as child word recognition data from Chapter 4 support the hypothesized geometry, separating labial and lingual features. Indeed, labial and coronal features are not regarded as mutually exclusive. Otherwise, coronal SR features would have caused a conflict with labial UR features in the contrast between /e/ and /ø/ in the MMN study, leading to a large MMN when the deviant was /e/. Also, such a conflict would stand in the way of learning words with coronal labial vowels, which was not the case in Chapter 4.

The results of the MMN study in Chapter 2 were the most conclusive regarding presence of asymmetries, hence, the remainder of my interpretation is mainly built on Chapter 2. As predicted, this MMN study showed an asymmetry between coronal and dorsal features in Dutch vowels /o/ and /ø/. This asymmetry is in line with underspecification of [CORONAL] and specification of [DORSAL]. This coronal-dorsal asymmetry replicates the findings for German vowels (Eulitz & Lahiri, 2004; Lahiri & Reetz, 2010; Scharinger et al., 2010), and is also attested in an MMN study on French vowels (De Jonge & Boersma, 2015). Current data supports underspecification of [CORONAL] and specification of [DORSAL] in the underlying representation of Dutch vowels.

This study is the first to show a perceptual asymmetry in adult native vowel perception between labial and nonlabial vowels in a language where labiality is lexically contrastive. Based on the hypothesized geometry and FUL's logic, I predicted that both directions of change in a labiality contrast would not result in a mismatch, and hence perception would be symmetrical, as indeed seemed to be the case in German. However, data for Dutch indicate that a labial vowel is not regarded an acceptable variant of a nonlabial vowel, but rather appears phonologically conflicting, as it yields an enhanced MMN response.

Although not previously attested in adults, a similar asymmetry has been found in infants. For six-month-old Danish infants, Polka and Bohn (2011) report an asymmetry between native nonlabial /e/ and labial /ø/ in the same direction as the labiality asymmetry I found in Dutch adults. However, Polka and Bohn report this asymmetry to no longer be present in 12-month-old Danish infants, indicating that experience with a vowel contrast may attenuate a perceptual asymmetry. Early infant speech perception does not necessarily reflect phonological processing yet. Adult asymmetries are not necessarily due to the same underlying mechanism as asymmetries in young infants.

Regarding the two-feature contrast between /e/ and /o/, a place asymmetry was predicted due to underspecification of coronal. However, a strong MMN response in both directions of change yielded no significant asymmetry and implied phonological conflict in either direction of change. A non-significant asymmetrical pattern was found which was in a similar direction as the labiality contrast: the labial vowel yielded a larger MMN amplitude. Findings for the two-feature contrast support a strong role for labiality in Dutch. This is different from German findings by Eulitz and Lahiri (2004). German findings show an asymmetrical pattern in the same direction as the place asymmetry would predict: an enhanced MMN when the deviant is coronal /e/. However, also in German, this trend did not reach significance, possibly due to the large acoustic distance between /e/ and /o/.

Overall, my predictions came true for the coronal-dorsal vowel contrast, where labiality was held constant. However, current data indicate a phonological conflict when a nonlabial vowel is changed into a labial vowel. FUL would predict that when a feature is not present, no mismatch can occur. Yet, the attested asymmetry implies that absence of a feature in UR does not render it impossible for a surface feature to yield a phonological conflict. In the next section, I propose that the feature geometry plays a role in how this might happen. I propose that (language specific) mapping rules define which SR features are mapped onto which UR features, and that this mapping employs the node structure. The hypothesized geometry and the mapping rules collectively:

(1) Ensure that the geometry reflects mutual exclusivity;
(2) Account for the labiality asymmetry we found and still explains the coronal-dorsal asymmetry;
(3) Provide an account for how a contrast between a feature and absence of an attribute – as is the case in LABIAL versus nonlabial – may be perceived.

Since Dutch and German may use different language specific mapping rules, the differences between the two languages may be explained within this model, assuming the same geometry for both languages. More detail is provided in Comparing Dutch and German.

**Feature geometry and mapping rules working together**

In order to deselect unwanted candidates and limit the number of word candidates, features extracted from the acoustics are compared to features stored in the underlying representation. Only the relevant comparisons are made. When, for instance, a feature [CORONAL] is extracted, only mapping to a mutually exclusive underlying feature like [DORSAL] would result in a meaningful/informative mismatch, whereas mapping to a feature like [HIGH] would not. The scope of the comparison of surface features must be defined to ascertain that the mapping process results in meaningful (no-)(mis)matches. FUL proposes a mapping procedure, but not all aspects are made sufficiently explicit.

I propose that the mapping procedure is guided by *mapping rules*. Which features are mutually exclusive is reflected by the feature geometry: nodes have mutually exclusive daughter features (this is also a principle assumed in FUL). Accordingly, mapping rules can use these nodes to define what should be compared to what. When a feature is extracted from the acoustics, this SR feature activates a certain rule that tells the phonological processing system what to do with the feature during the mapping process. In general, mutually exclusive features should be compared, so there could be a default mapping rule: an SR feature should be mapped to UR features below the same node that the SR feature belongs to. This way, meaningful matches, mismatches and no-mismatches occur, the latter strengthening activation of more probable candidates. This is the underlying assumption in FUL, as well.

For the current case of a labial-nonlabial vowel contrast, this implies that a labial SR feature is mapped to features below the LABIAL node. As a result, the mapping procedure may yield a match with a [LABIAL] UR feature, and it does not result in irrelevant mismatches with either [CORONAL] or [DORSAL]. However, when a [LABIAL] SR feature is compared to an underlying nonlabial vowel, the LABIAL node is not present in UR. Assuming the mapping rule guides phonological processing by pointing to the LABIAL node in UR as the location for meaningful mapping of features, the selection of the UR feature to which the surface feature is to be compared is unsuccessful in the case of a nonlabial underlying vowel. Consequently, the mapping procedure cannot be completed and no match, mismatch or no-mismatch is found. Rather, it can be regarded as an aborted mapping. Such termination of mapping implies that the underlying feature hierarchy is conflicting with the input. It follows that an aborted mapping can also be interpreted as a phonological conflict, similar to a mapping mismatch. Access to this conflicting form is prevented, leading to a large perceptual difference and hence large MMN response.

This explains the labiality asymmetry in the data: when a nonlabial vowel is mispronounced as a labial, this yields a phonological conflict, whereas a labial vowel mispronounced as nonlabial does not, because no labiality feature is extracted from this mispronunciation. Hence, no mapping of labiality features is triggered. In sum, assuming a role in mapping for the nodes in the geometry affects predictions for labiality as [LABIAL] is the only feature under the LABIAL node. A vowel is either [LABIAL] or lacks the LABIAL node altogether. Nonlabial is neither a phonological feature, nor does it have the stable acoustical features that would enable the perceptual system to extract it from the acoustics as a surface representation feature. When the acoustic feature corresponding to [LABIAL] is extracted from the signal and hence part of the surface representation, mapping it to the underlying representation with [LABIAL] results in a match. When the underlying representation lacks a LABIAL node, mapping cannot take place. This implies a phonological discrepancy between the surface and underlying representation, and we argue that such a case also indicates a phonological conflict. As such, this would predict an asymmetry in direction that was indeed found in the MMN study.

For the coronal-dorsal contrast, this does not change the predicted perceptual patterns. With this rule system, a coronal or dorsal SR feature is mapped onto the UR features underneath the LINGUAL node. Both [CORONAL] and [DORSAL] belong to this node. Assuming [CORONAL] to be underspecified and [DORSAL] to be specified, this mapping results in the asymmetries attested in Dutch as well as German.

The case of coronal/dorsal is different from the labiality contrast, because there are two mutually exclusive features which are contrastive, sharing their parental node. Also, rather than being underspecified (like CORONAL), nonlabial is not specified, and hence entirely absent from the representations. Underspecified coronal still has its parental node present, whereas nonlabial vowels have no labial node present. With a default mapping rule, mapping of coronal and dorsal can always be completed.



Mapping of features within the scope of their own node was already implied in FUL. However, more than a default rule appears imperative. Mapping articulator features at the level of labial or lingual nodes is suitable for Dutch vowels, because this way, labial is not mapped onto coronal or dorsal and meaningless mismatches are avoided. Yet, mapping restricted to either the labial or the lingual node is not suitable for Dutch consonants. Labial is mutually exclusive with coronal and dorsal in Dutch consonants. When following FUL in assuming unified consonantal and vowel features, a default mapping rule is not sufficient. For consonants, CORONAL, LABIAL and DORSAL are all mutually exclusive features, and hence should be compared to each other. This is not done with the default mapping rule which refers to the direct parental node. I propose that based on extracted feature combinations, adequate mapping rules can be triggered. For example, the combination of features [VOCALIC] and an articulator feature can elicit the default mapping rule, whereas a combination of [CONSONANTAL] with an articulator feature may elicit a specific mapping rule selecting the ARTICULATOR node as a whole. As a result, different rules may apply for consonants and vowels. Mapping of labiality may be language specific since it is not a universal contrast.

In sum, I have argued that for informative, and efficient feature mapping, rules of what should be compared to what are used. I proposed that these mapping rules define a scope of what feature mappings are relevant given the feature geometry. By doing so, both the labiality and place asymmetries found in Dutch adult vowel perception data can be accounted for. It can also explain how consonants and vowels can be treated differently in Dutch, still assuming unified features and featural organisation for both.

Furthermore, the coronal-dorsal asymmetry results from an asymmetry in mapping outcomes, whereas the labiality asymmetry is due to an aborted mapping in one direction of change. Different from FUL, not only features, but also the nodes in the geometry play a role during mapping. A consequence of this interpretation is that in addition to a mapping mismatch (as traditionally discussed in FUL), also an aborted mapping (when requirements of a triggered mapping rule are not met by the underlying representation) implies a phonological conflict in perception.

The combination of geometry and mapping rules can explain how phonological conflicts can be perceived between the feature labial and the absence of a labiality attribute in a nonlabial vowel. Perhaps this could be generalized to other contrasts. For instance, a similar asymmetry regarding the feature VOICE has been reported by Van der Feest and Fikkert (2015) in Dutch toddler's perception as well as in English in adult MMN studies (Hestvik & Durvasula, 2016). Van der Feest and Fikkert report that a change from a voiceless towards a voiced speech sound is perceptually more salient than vice versa, which resembles the case of labiality. A change from a non-feature (voiceless) towards a specified feature (voice) appears to result in phonological conflict whereas the reverse does not. As also pointed out by Lahiri (2018) in a footnote, [VOICE] could have its own node, similar to what I propose for [LABIAL]. Using a similar mapping rule, this would explain the voicing asymmetry, because mapping cannot be completed when a voiced sound is heard but mapped onto a voiceless UR, because the voicing node would then be absent, hence resulting in a phonological conflict. Whether this idea holds true should be assessed in future research, not only regarding voice and labiality, but also other potentially similar cases, like nasal, for example.

Additionally, the two adaptations may be fruitful in the field of automatic speech recognition (ASR) as well. For a computer algorithm modelling human speech perception to work well, one needs to explicate what should be compared to what: there is an acoustic form and a representation in memory that need to be compared in some way to get to lemma activation and word recognition. Perhaps the proposed mapping rules, which make use of feature geometry, would provide a worthwhile addition to current ASR algorithms.

**Alternative explanations**

Phonological theories vary from exemplar-based models, which assume all perceived exemplars to be stored, to abstractionist models, which assume one abstract underlying category for each phoneme. Abstractionist models may vary from fully specified to substantially underspecified representations (like FUL). Crucially, exemplar-based accounts assume all exemplars ever encountered need to be stored with all phonetic detail, whereas abstract accounts assume abstraction from phonetic detail. Exemplar-based theories come in different flavours. Some give a crucial role to frequency for explaining asymmetries, such as the Native Language Magnet

(NLM) theory (Kuhl, 1991; Kuhl et al., 2008), which predicts poorer discrimination from a more frequently occurring sound to a less frequent sound. Others rely on saliency for accounting for asymmetries, such as the Natural Referent Vowel framework (NRV) (e.g. Polka & Bohn, 2011; Masapollo, Polka, & Molnar, 2017), which bases predictions on focalization of formants (i.e. peripheral vowels are more salient than non-peripheral vowels).

An overview of the predictions discussed in Chapter 1 for exemplar-based accounts (frequency and saliency) and abstract accounts (full specification and underspecification) is given in Table 5.1. The experimental results for Dutch vowel perception are also presented.

**Table 5.1** Overview of predictions for different models.

| | | | Place | Labiality |
|---|---|---|---|---|
| **Exemplar-based** | **Frequency (NLM)** | | /ø/ → /o/ | /e/ ← /ø/ |
| | **Saliency (NRV)** | F1-F2 | /ø/ → /o/ | /e/ → /ø/ |
| | | F2-F3 | /ø/ ← /o/ | /e/ ← /ø/ |
| **Abstract** | **Full specification** | | /ø/ - /o/ | /e/ - /ø/ |
| | **FUL** | | /ø/ - /o/ | /e/ → /ø/ |
| | **Current interpretation** | | /ø/ ← /o/ | /e/ → /ø/ |
| | *Attested asymmetry* | | /ø/ ← /o/ | /e/ → /ø/ |

For each vowel pair and for each theory, the table indicates the direction of change in which the difference between vowels is perceptually largest. Marked in grey are predictions that are not in line with the experimental findings, which are given in the last line.

NLM predicts poorer discrimination in the direction from more frequent to less frequent. This model was built for accounting for within-category discrimination. If we expand it to between category discrimination, using the logic that frequent vowels accept more variation than infrequent vowels, since the prototype of a frequent vowel is a stronger magnet to other (deviating) vowels, it is expected that a change from a more frequent standard to a less frequent deviant would be more difficult to discriminate and hence show a smaller MMN effect than vice versa. Dutch front round vowels like /ø/ are relatively infrequent compared to front (unround) vowels and back (round) vowels. Thus, discrimination would be easier when the deviant is not a front round vowel (so if the deviant is /e/ or /o/). These predictions are opposite to the findings in the current study.

Phonetic saliency has been used to account for asymmetries in vowel perception among others by proponents of the Natural Referent Vowel (NRV) framework (Polka & Bohn, 2011, Masapollo, Polka, & Molnar, 2017; Masapollo, Polka, & Ménard, 2017). It assumes that vowels with formant frequencies closer together have more focalized energy, and hence, universally are more salient in perception than vowels with formants further apart. Changes from less to more focal vowels are assumed easier to discriminate. Predictions differ as to whether one takes F1 and F2 or F2 and F3 to determine focalization. Based on closeness of F2 and F3, /e/ is more focal than /ø/, although toward the end of the vowel the difference becomes small, and the formants are slightly overlapping, making them less distinguishable. Based on the F1-F2 dimension, /ø/ would be more focal than /e/, but the formants do not get close, and it remains the question as to whether the formants lead to focalized energy. Moreover, this latter prediction is contradicting the predictions made in Polka and Bohn (2011). It must be noted that the results reported in Polka and Bohn (2011) do not always conform to their own predictions. Notably, they report better discrimination in the direction from /e/ to /ø/ in Danish infants.

In other words, the predictions based on the NRV are not entirely straightforward. For /o/-/ø/ the predictions are not clear either. Based on F1-F2, /o/ is more focal than /ø/. Based on F2-F3 /ø/ is more focal than /o/. NRV is not explicit on the relative weighting of F1-F2 focalization and F2-F3 focalization. For the place contrast, predictions are in line with the results if predictions are based on F2-F3, but not when basing predictions on F1-F2. For the labiality contrast, it is the other way around: predictions hold true for the F1-F2 dimension but not for the F2-F3 dimension.

The asymmetries revealed in the current study cannot easily be explained by the two exemplar accounts discussed. The current NRV framework only partly predicts the attested asymmetries. The frequency account makes predictions opposite to current experimental finding. Abstract but full specification accounts using bivalent features would predict symmetrical perception in all vowel contrasts. Regardless of the direction of change, feature values would be conflicting. Since abstract full specification does not predict asymmetries at all, current findings are not in line with those models either.

The current study shows evidence for asymmetrical processing of vowels and the attested coronal-dorsal asymmetry provides support for the FUL model. However, the FUL model does not make clear predictions with respect to the labiality contrast since it assumes labial to share its direct parental node with coronal and dorsal in FUL's geometry. It is not clear how labial and lingual features are separated. The currently attested labiality asymmetry cannot be explained within the traditional FUL model, since findings imply a phonological conflict in a case where the UR is not specified. In FUL, something that is not specified would accept variance.

Solving the issue of mutual exclusivity with a different geometry than FUL uses, I still predicted a coronal-dorsal asymmetry, which is in line with current findings

The labiality asymmetry I found can be explained if in addition to the split between labial and lingual features in the geometry, mapping rules that use the nodes in the geometry are assumed.

This interpretation shows how listeners might evaluate incoming signal on its phonology. Attested asymmetries confirm the predictions of this interpretation.

## Implications for the Dutch vowel system

The current section discusses the implications of the two proposed adaptations to the FUL model for the Dutch vowel system. I will only make claims regarding the specification of labial and lingual features. As an answer to the research question how front (unround), front round and back (round) vowels are phonologically represented in a language where both place and labiality contrasts are lexically contrastive, I present the following surface and underlying representations for Dutch vowels /e/, /ø/ and /o/. From the feature tree, only the ARTICULATOR node is depicted.



Surface representations:

```
    /e/              /ø/                 /o/
     |                |                   |
    ART              ART                 ART
     |               /   \               /   \
  LINGUAL       LABIAL   LINGUAL     LABIAL   LINGUAL
     |             |        |          |         |
 [CORONAL]     [LABIAL] [CORONAL]   [LABIAL] [DORSAL]
```

Underlying representations:

```
    /e/              /ø/                 /o/
     |                |                   |
    ART              ART                 ART
     |               /   \               /   \
  LINGUAL       LABIAL   LINGUAL     LABIAL   LINGUAL
     |             |        |          |         |
    [Φ]        [LABIAL]    [Φ]      [LABIAL] [DORSAL]
```

Labiality and place contrasts in tense mid vowels received most attention in the current thesis, but similar contrasts exist for high vowels /i, y, u/ and the lax vowels /ɛ, ɔ, ʏ/. I assume similar representations for the labial and lingual features for these high vowels.

I propose that the LINGUAL and LABIAL nodes are part of the representations of Dutch vowels, and they are used during the mapping procedure to navigate which SR feature is compared to which UR feature. In combination with underspecification of CORONAL, specification of DORSAL, specification of LABIAL, and no specification for nonlabial, this provides an explanation for the current Dutch vowel perception data.

**Dorsal and labial redundancy**

In Dutch vowels, a feature [DORSAL] implies a feature [LABIAL], but not vice versa. Redundant features may be underspecified in FUL. Thus, it is conceivable that [DORSAL] and [LABIAL] need not both be specified. Perhaps [LABIAL] could be specified for coronal vowels, but underspecified for dorsal vowels. First, I will consider what this would imply for perception. Second, I will provide a developmental perspective.

If [LABIAL] is underspecified for the dorsal vowel /o/, this would not change the predicted perceptual outcomes for the contrast /o-ø/. A change from /o/ to [ø] would result in a phonological conflict, because the coronal surface feature of /ø/ mismatches with the dorsal UR feature of /o/. The labial SR feature of /ø/ results in a no-mismatch with the underlying representation of /o/ if it is underspecified for labial. After all, if /o/ is underspecified for labial, it still has both a LABIAL and a LINGUAL node, just like when [LABIAL] would be specified. Vice versa, when /ø/ changes into [o], labial features match, and place features result in a no-mismatch, so there is no phonological conflict in this direction of change. Thus, a larger perceptual difference is predicted when /o/ changes into [ø] than vice versa, which is in line with the perceptual asymmetry found for this vowel contrast in the current research.

Labial underspecification in dorsal vowels would not imply different predictions for vowel contrast /o-e/ either. If /o/ changes into [e], the surface feature coronal of [e] would mismatch with the underlying dorsal feature of /o/, hence resulting in phonological conflict. No labiality mapping is initiated in this direction of change. Vice versa, when /e/ changes into [o], mapping of place features results in a no-mismatch, since /e/ is underspecified for CORONAL. A phonological conflict is also predicted regarding labiality, since a [LABIAL] surface feature of [o] cannot be mapped to the representation of /e/, since /e/ has no LABIAL node. Ultimately, a phonological conflict is predicted in both directions of change for /o-e/, which is similar to when [LABIAL] would be stored for /o/.

To conclude, predictions for perceptual outcomes would not be different when LABIAL is either specified or underspecified for dorsal vowels. Perceptual data is in line with both accounts. Perhaps the acquisition process can enlighten us to what would be the most probable position.

Similar to what has been proposed by Levelt (1994), and Levelt and Fikkert (2008) among others, Lahiri (2018) assumes that in acquisition, the contrast between LABIAL and nonlabial (coronal by default) vowels is established before a coronal-dorsal contrast is established. Thus, initially, dorsal may be redundant with [LABIAL], and dorsal is probably not specified. At a later stage of acquisition, Dutch children add coronal labial vowels to their vowel inventory. Dorsal consonants will by then have entered the child's phonology. These changes indicate (a) that dorsal is no longer redundant, since place and labiality do not coincide anymore, and (b) that a coronal-dorsal contrast is important at least in consonants. If children start with a labial vs. lingual contrast where CORONAL is the default lingual feature, [DORSAL] is required to be added to this initial setup for labiality and place to become mutually autonomous contrasts. If the system starts with a specified feature [LABIAL], and [DORSAL] is added later, this would imply both LABIAL and DORSAL to be specified for back vowels. Although this seems most plausible, it is not precluded that stored features can become underspecified features at a later stage of acquisition.

5

In summary, although [LABIAL] appears redundant for back vowels in Dutch, current data cannot enlighten us regarding whether or not labial is underspecified in dorsal vowels. However, from an acquisition perspective, it seems plausible that despite the fact that dorsal vowels are always labial in Dutch, dorsal and labial features may both be specified.

**Lax mid vowels**

In Chapter 3, the contrasts between lax mid vowels /ɛ, ʏ, ɔ/ were explored in a semantic priming experiment. As discussed in Chapter 2 and 3, Dutch tense mid vowels tend to diphthongize slightly and may end as high vowels – termed semi-diphthongized vowels in Adank et al. (2004) – whereas lax vowels are true monophthongs. Since German tense vowels are not diphthongized, one might argue that this is a reason for the difference between German and Dutch. However, as already discussed in the experimental chapters, predicted outcomes with respect to labiality and place contrasts are similar regardless of the vowels being diphthongized or not, since the shift merely strengthens place features. Moreover, data collected in Chapter 3 did not reveal different outcomes for tense and lax vowels.

I assumed that lax mid vowels /ɛ, ʏ, ɔ/ have similar representations regarding place and labiality as tense mid vowels /e, ø, o/, and hence no difference between tense and lax vowels was predicted with respect to (a) symmetrical perception of contrasts. However, whether or not the three lax vowels are the same height is not entirely clear. In Chapter 1 we have seen differences in the literature when it comes to height of these vowels. Particularly one of the lax vowels is disputed; some may argue that /ʏ/ is a high vowel rather than a mid vowel (Rietveld et al., 2004), whereas others consider it mid on the same level as /ɛ/ and /ɔ/ (e.g. Moulton, 1962; Zwaardemaker & Eijkman, 1928). Booij (1989) also considers /ʏ/ to be a mid vowel, but he assumes four height levels and considers /ʏ/ to be  high mid, whereas /ɛ/ and /ɔ/ are considered low mid. It is hard to resolve this lack of consensus, since /ʏ/ can be distinguished from all other vowels either as a high or as a mid vowel.

If /ʏ/ is a high vowel, /ɛ/ and /ɔ/ can be distinguished merely based on either labiality or place, since in that case labiality and place coincide for the lax mid vowels. This could be similar to the early stage in acquisition as discussed above. Since the contrast between /ɛ/ and /ɔ/ yielded inconclusive results in Chapter 3, we cannot draw conclusions on whether or not this contrast is represented in the same way as their tense counterparts.

Furthermore, if /ʏ/ is high rather than mid, perhaps perceptual asymmetries for the contrasts between /ʏ-ɛ/ or /ʏ-ɔ/ could be based on height, similar to the height asymmetries in French (De Jonge & Boersma, 2015) and English (Scharinger et al., 2012). These studies assume MID to be underspecified. This would predict that /ɛ/ and /ɔ/ would accept /ʏ/ as a variant based on height, whereas /ʏ/ would not accept mid vowels /ɛ/ and /ɔ/ as surface variants. Another possible account is that rather than being underspecified, MID is not specified at all (suggested by Lahiri, 2018). Then, in FUL's logic, mid vowels consequently lack a HEIGHT node. This would give opposite predictions compared to when MID is underspecified if I assume that this leads to phonological conflict when one attempts to map a high SR vowel onto a mid UR vowel, due to an aborted mapping.

So based on height, there are two possibilities: mid is underspecified or mid is not specified. These two options provide opposite predictions of what direction of change would yield a phonological conflict based on height. This implies that either direction of an asymmetry for these contrasts could be explained based on height.

In sum, it depends on /ʏ/ being HIGH or MID whether or not the representations of labiality and place for lax /ɛ/ and /ɔ/ are similar to their tense counterparts /e/ and /o/. Our data cannot enlighten us regarding the status of /ʏ/ or whether mid vowels are underspecified or not specified for height. This needs further research.

### Comparing Dutch and German

Asymmetries in the perception of coronal versus dorsal vowels have been attested for both Dutch and German listeners. An asymmetry between labial and nonlabial vowels has been found for Dutch, but has not been reported for German (compare Eulitz and Lahiri, 2004: German; Chapter 2: Dutch). I have provided an account regarding perception and representation of the labiality contrast in Dutch, which could explain the attested perceptual asymmetry. Still, the question remains as to why perception of this contrast would be different in German.

Eulitz and Lahiri (2004) reported symmetrical perception for the vowel pair /e/-/ø/. These vowels share place of articulation - they are both coronal - and hence this contrast does not constitute a conflicting situation with respect to place of articulation. Absence of an asymmetry for this contrast was replicated for German listeners in an MMN passive oddball paradigm using both words and nonwords (Cornell et al., 2011). This is unlike the results in the current dissertation, which reports an asymmetry for labiality in Dutch.

At first glance, German and Dutch vowel systems seem rather similar regarding their vowel inventories. Like German, Dutch has a three-way contrast between /e/, /ø/, and /o/. Yet, looking at the linguistic system as a whole, the nature of the front round vowels appears different in Dutch and German.



In monomorphemic (underived) context, front round vowels are far more frequent in Dutch than in German. A count in CELEX (Baayen et al., 1995) accessed through Reelex Version 0.4.5 (Reetz – Celex interface) shows there are virtually no minimal pairs in German monomorphemic monosyllabic nouns with a labial vs. nonlabial front vowel contrast. For tense vowels, there are nine monosyllabic monomorphemic German nouns with /ø/ (of which only three occur more than once in the database), and only two monomorphemic monosyllabic nouns with /y/. There is only a single labial vs. nonlabial minimal pair: T[y]r (*door*) vs. T[i]r (*animal*). There are many more words with coronal labial vowels in Dutch (76 with /ø/ and 37 with /y/), with at least 13 labial vs. nonlabial minimal pairs for the high vowels (only nouns), and at least 8 minimal pairs words with mid vowels (only nouns).

Since Dutch has several minimal pairs and labial coronal vowels occur rather often in monomorphemic words, distinguishing between labial and nonlabial vowels at a phonological level is key to resolving confusion. In Dutch, a [LABIAL] feature should trigger mapping of labiality features. This is not required in German. If two German word stems differ with respect to labiality of a vowel, they typically differ regarding more phonemes than just that vowel. Hence, the difference between the words also can be made based on other featural differences than labiality of the vowel. Unless, of course, in the case of minimal pairs, which are rare in German. In those cases, context aids in resolving confusion. Both options T[y]r (*door*) and T[i]r (*animal*) will not be equally plausible in a given sentence. Accordingly, the labial versus nonlabial distinction in German does not require mapping of labiality features in order to learn or recognize all words in the language.

If no mapping of labiality features takes place in German, a [LABIAL] surface feature does not result in conflict with a nonlabial underlying form. This way, no conflict would occur in either direction of change between a labial and a nonlabial vowel in German, as is in line with the symmetry reported by Eulitz and Lahiri (2004) and Cornell et al (2011) for the /e/-/ø/ contrast in German. Dutch words with front round vowels are true phonemes, and form minimal pairs with other words, whereas in German they are often allophones. Although /ø/ occasionally also occurs in nonderived (lexical) environment, its contrastive value in German is limited in comparison to Dutch, which may explain why German and Dutch listeners react differently to [LABIAL] in the context of [CORONAL] vowels: in German [LABIAL] is not a strong lexical contrast, while in Dutch it is.

However, in morphologically complex words front round vowels are more frequent in German than in Dutch. Front round German vowels are typically derived from back vowels, due to morphological umlaut. For example, [ø] often arises as the result of morphological umlaut as in the plural in *V[o]gel* – *V[ø]gel* ('bird(s)'), and in the comparative form of the adjective in *hoh* [o] – *höher* [ø] ('high – higher'). Thus, [ø] often occurs in derived environments, rather than in the lexicon. Modern Dutch, however, does not have morphological umlaut. Dutch hardly has alternating stems, and non-alternating stems with front round vowels are the norm in Dutch.

An alternative approach to the representation of morphologically related words is presented in Scharinger (2009), and Scharinger et al. (2010). The authors assumes that dorsal vowels in lexical words that alternate in derived forms (i.e. certain morphological categories), such as the German word *Vogel*, have a different underlying representation than words with dorsal vowels that do not alternate. Specifically, they argue that the /o/ in *Vogel* is underspecified for [DORSAL], but specified for [LABIAL]. Under this view [ø] does not mismatch with [o] in *Vogel*: the perceived [CORONAL] no longer mismatches, as [DORSAL] is not part of the underlying representations. /ø/ and $/o/_{alternating}$ have the same underlying representation, i.e., [LABIAL], but at a later stage a default coronal fill-in rule and a specific dorsal fill-in rule apply based on morphology, differentiating both vowels in the surface representation.

In terms of the feature hierarchy proposed in this thesis, this would mean that these alternating vowels, which are underspecified for [DORSAL], lack specification under the LINGUAL node in the underlying representation. A separation of labial and lingual features in combination with mapping rules as proposed in this thesis would not change predictions for discrimination of coronal and dorsal vowels compared to Scharinger's predictions for perception in German alternating words. This is because the LINGUAL node would still be present when [DORSAL] is underspecified. Thus, mapping of coronal and dorsal features remains similar to the traditional way following FUL as proposed by Scharinger (2009) and Scharinger et al. (2010). German findings regarding place both in alternating and non-alternating forms are compatible with the current interpretation. However, since Dutch does not have similar vowel alternations like morphological umlaut in German, there is no reason to assume that Scharinger's alternative approach applies in Dutch, i.e. underspecification of [DORSAL] seems irrelevant in Dutch vowels.

The role of labial, appears different in the linguistic systems of Dutch and German. In terms of the mapping rules proposed in the current thesis, German data implies that [LABIAL] does not trigger mapping of labiality features, as perception of the contrast between /e/ and /ø/ was symmetrical. However, the difference between nonlabial /e/ on the one hand and labial vowels /ø/ and /o/ on the other hand (the latter two being related due to umlaut) is important in German, as /e/ should not provide access to a dorsal vowel, neither in alternating or non-alternating stems. Rather than mapping labiality features, it appears that a perceived [LABIAL] will automatically activate back vowels in German. In that sense, [LABIAL] is used to mark the contrast between /e/ and /ø/. Hearing a front round vowel in German immediately activates the morphology. While testing this is beyond the scope of this dissertation, it certainly warrants further research. There is some evidence in the literature that listeners are sensitive to critical phonological properties in parsing morpho-phonological forms (e.g. Post, Marslen-Wilson, Randall & Tyler, 2008, Pliatsikas, Wheeldon, Lahiri, & Hansen, 2014).



In German, the distinction between /e/ and /o/ - which differ regarding both place and labiality - could be solely based on place, predicting a place asymmetry, but no labiality asymmetry, which is what Eulitz and Lahiri (2004) predicted. Indeed, perception of this vowel pair in German showed a larger MMN response when the deviant was coronal. Although this was not a significant asymmetry, it was still in the same direction as the coronal-dorsal asymmetry found for single-feature contrast /o/-/ø/. The lack of significance for /o/-/e/ may be due to the larger acoustic distance for this contrast compared to /o/-/ø/ (Eulitz & Lahiri, 2004).

If, in German, the perceptual distinction between /e/ and /o/ is based merely on place since labiality is not mapped, this would imply that /e/ may prime /o/ when /o/ is part of an alternating word stem, and therefore underspecified for [DORSAL]. As far as I know, this has not been investigated yet, but would be an interesting suggestion for future research in semantic priming research.

For Dutch, a different prediction should be made for /e/-/o/. For place, the same asymmetry is expected as for German, but for labiality, the two languages differ. Since both labiality and place can result in phonological conflicts in Dutch based on mapping of features, both directions of change will be conflicting for the contrast between /e/ and /o/. Based on place, I would expect an enhanced MMN in the context [e]/o/, but not vice versa. However, based on labiality, an enhanced MMN in the context [o]/e/ would be predicted. One explanation for the lack of a significant asymmetry may be that these cancel each other out. Another possible explanation is that the two vowels are acoustically very different and easy to discriminate, similar to how Eulitz and Lahiri explain the absence of a significant asymmetry for /e/-/o/ in German.

However, Scharinger et al. (2016) suggested that with featural distances larger than one feature, the underspecification interpretation based on FUL's mapping procedure may not hold, illustrating the benefit of phonological information for processing acoustically minimal sound differences. In the fMRI study by Scharinger et al. (2016) it was observed that in vowels differing by two or more features, a change from a less specified sound to a more specified sound elicited stronger responses in German listeners than vice versa. This does not explain the pattern in German /e/-/o/ found by Eulitz and Lahiri (2004), which is in the same direction as a coronal-dorsal asymmetry, merely nonsignificant, but the Dutch pattern is in opposite direction from what would be predicted based on place only, and one could argue that this is similar to what happened in the study by Scharinger and colleagues (2016). However, what causes the opposite direction of the asymmetries is not entirely clear. Possibly, this is similar to the explanation of acoustic distance provided by Eulitz and Lahiri (2004) in the sense that acoustics plays a larger role when acoustic distance is larger, whereas phonological information is particularly important when acoustic distance is smaller. Although this may explain absence of an asymmetry, it is not entirely clear why this would result in an opposite asymmetry. It goes beyond the scope of this dissertation, but further investigation of asymmetries in multiple-feature contrasts versus single-feature contrasts would be relevant in future research to gain better understanding of this matter.

In summary, I assume the same node structure for Dutch and German. I also assume similar representations for the lingual features as well as similar mapping rules in both languages for lingual features. Also, because of the different role labiality plays in German and Dutch, I propose that the labiality contrast does not require mapping of labiality features in German, whereas this is required in Dutch. Thus, the difference in the perceptual data for the labiality contrast in German and Dutch may be explained by assuming a difference regarding feature mapping rules.

### Comparing children and adults

The three-way distinction between back (round), front (unround) and front round vowels (i.e. for mid vowels /o, ø, e/) in adult phonology may be described in multiple ways, as has been proposed in previous literature (see Chapter 1). Perhaps insight into how children build up the system of vowel contrasts may provide insight into the adult system?

A three-way contrast between [CORONAL], [LABIAL] and [DORSAL] is established earliest in consonants. To form such a three-way contrast, the feature dorsal may be added to the initial labial-coronal contrast. For consonants, [LABIAL] is contrastive with both [CORONAL] and [DORSAL]. However, for Dutch vowels it is not sufficient to simply add DORSAL to account for the three-way distinction between front, front round and back (round) vowels, since in Dutch vowels [LABIAL] is not mutually exclusive with [CORONAL] or [DORSAL]. Rather, [LABIAL] may be combined with either of those place features.

In the early phase of acquisition, it may be sufficient to compare LABIAL to LINGUAL (hence by default [CORONAL]) to distinguish both consonants and vowels. This is the opposition established by the ARTICULATOR node if it splits into LABIAL and LINGUAL. At this stage, merely the coronal lingual feature may be part of the phonology, as it is regarded default. After DORSAL is added, vowel feature mapping in perception needs to occur for lingual and labial features separately. [LABIAL] would otherwise mismatch with [CORONAL] or [DORSAL] features, while it should not. For consonants, labial and coronal remain mutually exclusive.-

It is unclear, however, when or how the children have arrived at this stage in acquisition. Data from Van der Feest and Fikkert (2015) showed that Dutch two-year-olds are sensitive to /t,k/ mispronunciations in a syllable if the vowel is back, but not when the vowel is front like /i/. Clearly, separation of vowels and consonants in word representations is not yet fully established at this stage. To track the exact course of development/acquisition from a labial-coronal contrast to the adult system which handles both vowels and consonants in different ways in Dutch, future research is required to reveal a more precise time course of how the Dutch vowel system is built and what the exact order of acquisition is for both feature representations and mapping rules.

As a first step in finding out more about the age at which children acquire the Dutch vowel contrasts, I tested three-year-olds, because this is the first age when front round vowels are produced (Beers, 1995). Thus, representations for front round vowels have been formed. However, representations of front, front round and back (round) vowels may not yet be as stable as in adults. The intermodal preferential looking experiment in Chapter 4 showed that, similar to adults, Dutch three-year-olds are sensitive to labiality mispronunciations as well as place (lingual) mispronunciations. All mispronunciations, irrespective of the direction of change, hindered word recognition. This indicates that the toddlers in this study were sensitive to the three-way distinction of Dutch tense mid vowels.

The fact that correctly pronounced words (CPs) with /ø/ were successfully recognized implies that, just like adults, Dutch three-year-olds no longer consider [LABIAL] and [CORONAL] to be mutually exclusive. Otherwise, [CORONAL] and [LABIAL] would be mapped onto each other in perception just like in the early stage of word representations. The [CORONAL] surface feature of [ø] would then mismatch with the [LABIAL] underlying feature of /ø/. If so, word recognition would be hindered for correctly pronounced words with /ø/. Yet, children successfully recognized the target words in this condition. That word recognition is not hindered for CPs with /ø/, implies that LABIAL and LINGUAL features are kept separate during feature mapping.

In sum, toddlers show a rather adult-like phonology: the three-year-olds discriminate all tested vowel contrasts between tense mid vowels, and recognize all CPs. Nevertheless, I cannot conclude toddlers' and adults' processing of the contrasts to be identical. After all, adult data showed asymmetries, whereas no asymmetries were found in toddlers' perception data.

For the place contrast, adult data showed a coronal-dorsal asymmetry, supporting underspecification of [CORONAL]. No such asymmetry has been revealed in child data. This does not necessarily mean that toddler perception was indeed symmetrical. As discussed in Chapter 4, IPLP may have been not sensitive enough to reveal asymmetries for this contrast in this age group. Perhaps children were already too proficient to reveal any asymmetries through this paradigm?

For the labiality contrast, the proportion of target looking was negative when /e/ was mispronounced as [ø]. This may indicate that the labial mispronunciation of nonlabial /e/ was not accepted as a potential variant, and the MP may in fact be interpreted to be a different word, hence referring to the distractor rather than the target item. Vice versa, when /ø/ is mispronounced as [e], no negative PTL was found. This was not a significant asymmetry, but this trend is in line with the labiality asymmetry revealed for adults (Chapter 2) for the same contrast.

The question is whether this asymmetrical trend in toddlers occurs for the same underlying reasons as the asymmetry in adults. If it is due to an aborted mapping of the [LABIAL] SR feature due to absence of a LABIAL node in UR, the same patterns would be expected to occur for the two-feature contrast /o-e/, since they also differ regarding labiality: a labial (SR) should conflict with a nonlabial (UR). We did not intend to find asymmetries between different types of conflict, so sample size may have been too small to reveal such a difference. However, the same pattern we found in the MMN experiment with adults was present in the child data; in neither of the directions the mispronunciation led to recognition of the target word, and when /e/ was mispronounced as [o], a significant negative naming effect was found. This indicates that the MP is interpreted to label the distractor item. This is in line with the idea that mapping of the [LABIAL] SR feature is impossible in this direction of change, due to an absence of a labiality node in the UR of the CP. This indicates phonological conflict between SR and UR, leading to the CP not being considered a viable candidate for the MP.

To conclude, child data appears adult-like, although no strong conclusions can be drawn since child data did not show significant asymmetries. Still, both children and adults show asymmetrical patterns in the same direction for the labiality contrast (one-feature as well as two-feature). Toddler data is in line with the notion that labial features do not trigger mapping to lingual features, but do trigger mapping to labial features, which is not possible when the UR does not have a LABIAL node. Future research should track development of the vowel system throughout acquisition in younger children, to investigate how the system is gradually built.

## Limitations and suggestions for future research

In the current thesis, I investigated how place and labiality features of vowels are represented in Dutch, focusing on the three tense mid vowels /e, ø, o/. For this set of vowels, I collected perceptual evidence using different methods and testing different populations. Mismatch Negativity responses to vowel changes indicated asymmetrical perception for Dutch adult listeners in Chapter 2 for both labiality and place of articulation. This was, however, not confirmed in the other chapters.



In Chapter 3, designing a good semantic priming experiment proved to be challenging in a language with a limited number of lexical items that meet the rather strict criteria for stimuli. For this reason, semantic priming appeared unsuitable to test the vowel contrasts of interest in Dutch words. For future research I would suggest the following alternatives.

Using EEG, one could test perception of phonological contrasts in words in an MMN study using words or word parts, similar to what has been done in German (Scharinger et al., 2010). This is not a radically different method compared to the MMN study on vowels in isolation conducted in Chapter 2. In order to see whether results can be replicated using a more fundamentally different method one could also measure N400 during lexical decision of words and pseudowords (mispronunciations) similar to Friedrich, Eulitz, and Lahiri (2006). The N400 is suggested to be sensitive to the time-course of cognitive processes underlying word recognition. The N400 has not only been correlated to aspects of semantic processing in sentence and priming contexts, but also to single word processing (for a review, see Kutas & Federmeier, 2000). Hence, this method does not require semantically related word pairs, and suitable test items will be much easier to select.

Alternatively, one could measure reaction times rather than electrical brain activity. Lahiri and Reetz (2010) report on a speeded auditory lexical decision task in German. The rationale in this paradigm is that when mapping of features of a mispronunciation does not result in a phonological conflict (e.g. a mismatch), it is more difficult to reject the MP as a nonword than when mapping does yield a conflict. Hence, it is harder (i.e. takes longer) to decide whether the word is a lexical item. In this paradigm, mispronunciations of words are used, but no semantically related lexical items are required. Consequently, there are fewer criteria to take into account for selecting appropriate stimuli and the pool of potential stimuli would be larger. A speeded auditory lexical decision task could thus be a promising paradigm to test labiality and place vowel contrasts in Dutch words.

In Chapter 4, the Intermodal Preferential Looking experiment showed that Dutch three-year-olds were probably too proficient for this method to be sensitive enough to reveal perceptual asymmetries. In other words, the task may have been too easy, and children paid attention to the details of pronunciation just like what would happen if adults would perform a similar task in easy listening conditions. It is therefore difficult to say whether the toddlers use the same phonology as adults do, nor can I make elaborate claims about how the adult system is built. To answer these questions, future research should track development throughout acquisition. I would suggest to complement this study with studies in younger children, developing from phonetic to phonological listeners. Around the age of 6 months, infants have tuned into their native language with respect to vowels. vowels (Kuhl et al., 1992; Tsuji & Cristia, 2014). Thus, one could test infants' perception prior to this, for instance at the age of 4 months, as well as after this age, for instance at the age of 8 months. One could run an MMN study in for example 4- and 8-month-olds, similar to what we conducted in adults in Chapter 2, to see what changes over the course of this important period in development. In slightly older childrenwho have developed a lexicon  (e.g. 18 months), testing vowel discrimination using for instance IPLP might provide more insight into what the impact of the developing lexicon is, as at this point phonological processing becomes relevant.

For the currently tested age group, an alternative to IPLP would be to use a more sensitive method such as EEG (for example MMN or N400 experiments as discussed above), which measure electrical brain activity during processing rather than overt behaviour. The MMN study described in Chapter 2 (adults) could be performed with three-year-olds in slightly adapted form, to assess whether or not they show similar asymmetries to adults with this more sensitive method. For younger children, IPLP may be sensitive enough to reveal asymmetries.

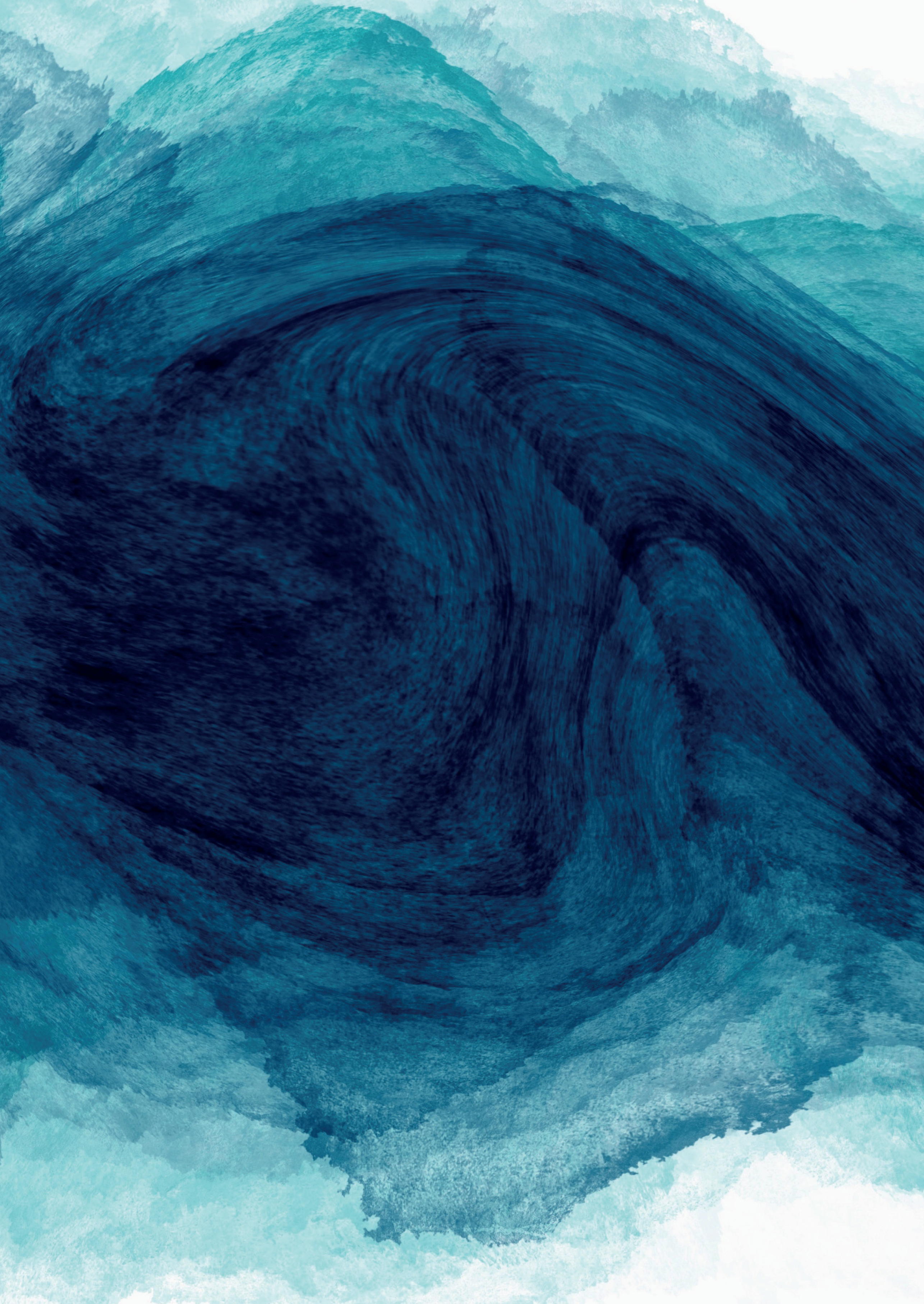
## Conclusion

Despite the limitations of some of these experiments, I have shown some important findings.

First, my experiments provide evidence for asymmetries in vowel perception in Dutch for both labiality and place contrasts. This provides support for an abstract underspecification account.

Second, there may be more ways in perception for a phonological conflict to occur. Apart from phonological conflict in perception due to a mapping mismatch (as relevant for the coronal-dorsal vowel contrast), phonological conflict can also be due to an aborted mapping - as relevant for the contrast between labial and nonlabial vowels in Dutch. If indeed this different type of phonological contrast occurs in cases like the labiality contrast, I do not only expect this to be replicated in Dutch at word level, but future perception experiments might show whether this is similar in other languages where labiality is a lexical contrast, such as in French or in Scandinavian languages. Future experiments could also show whether the notion of an aborted mapping replicates to similar types of contrasts between a feature and no feature (e.g. voicing). Also, longitudinal studies tracking language acquisition may provide greater insight in how phonological systems are built during development regarding place and labiality contrasts in vowels.

Lastly, Dutch and German may seem similar on a superficial level when considering their vowel inventories, but they behave differently. Differences in the language system may impact what is stored or underspecified, or which mapping rules are required. Ultimately, even when phonological contrasts in different languages appear similar, they may be different below the surface. Detailed cross-linguistic studies assessing phonological contrasts would be valuable in providing more insight into similarities or differences between language systems, and consequently also into language general or language specific phonological phenomena. An eye for detail portrays the bigger picture.



# Chapter 6

## Samenvatting in het Nederlands

## Nederlandse klinkers onderscheiden

Om te begrijpen wat iemand zegt, is het belangrijk de verschillende woorden die iemand zegt goed te herkennen. Woorden verschillen soms slechts wat betreft een enkele spraakklank, maar hebben een volstrekt verschillende betekenis. De woorden 'boer' (uitspraak: /bur/) en 'buur' (uitspraak: /byr/) verschillen bijvoorbeeld slechts wat betreft één klinker. Het verschil tussen de klank /u/ (zoals in 'boer') en de klank /y/ (zoals in 'buur') wordt veroorzaakt door een verschil in tongpositie. De /u/ is een achterklinker, waarbij de tong een constrictie vormt achter in de mond. De /y/ is een voorklinker, waarbij de voorkant van de tong omhoog komt. Hier zijn 'voor' en 'achter' fonologische kenmerken van de klinkers. Het onderscheiden van de verschillende kenmerken van klanken is essentieel voor het juist herkennen van woorden wanneer je naar spraak luistert. Je brein doet dit zonder dat je je er bewust van bent en ook nog eens ontzettend snel tijdens elk gesprek dat je met iemand voert. Hoe doen luisteraars dit?

Het onderscheiden en herkennen van fonologische kenmerken is geen vanzelfsprekendheid. Elke keer dat je hetzelfde woord hoort, wordt het telkens een beetje anders uitgesproken. Zelfs wanneer dezelfde persoon een woord meerdere keren herhaalt, zal het nooit tweemaal exact hetzelfde worden uitgesproken en dus ook niet exact hetzelfde klinken. Behalve tussen verschillende woorden zijn er dus ook verschillen tussen exemplaren van dezelfde woorden. Zulke verschillen zijn niet betekenisvol: het gaat nog steeds om hetzelfde woord, bijvoorbeeld twee varianten van het woord 'boer'. Je brein moet dus weten welke verschillen belangrijk genoeg zijn om opgemerkt te worden en welke niet. De verschillen die zorgen voor betekenisverschillen zijn belangrijk, dus de verschillen tussen verschillende woorden. Deze belangrijke verschillen worden fonologische contrasten genoemd. Je brein dient gevoelig te zijn voor deze verschillen, maar andere verschillen die niet voor een betekenisonderscheid zorgen, kunnen worden genegeerd.

Om woorden te herkennen, dien je als luisteraar wat je hoort (de akoestische input) te vergelijken met de woorden zoals die in je brein zijn opgeslagen. Deze opgeslagen vorm van woorden wordt de mentale representatie genoemd. Een belangrijke kwestie binnen de fonologie is welke informatie in deze representaties is opgeslagen. Hoe stellen deze representaties ons in staat op efficiënte wijze woorden te herkennen? Hoe worden belangrijke verschillen opgemerkt, maar betekenisloze verschillen genegeerd? Om hier meer inzicht in te geven, doe ik in dit proefschrift verlag van mijn onderzoek naar de mentale representaties van Nederlandse klinkers.

Verschillen tussen klinkers worden met name veroorzaakt door verschillen in tongpositie en/ of lippositie tijdens de articulatie. Er kunnen daarmee op drie manieren verschillen worden gemaakt:

1. **Lipronding.** De lippen kunnen gerond worden (zoals bij de /o/ in 'boot') of zij worden niet gerond (zoals bij de /e/ in 'been').
1. **Plaats.** De tong kan aan de voorkant of aan de achterkant omhoog komen en dus een voor- of achterklinker vormen, zoals in het voorbeeld 'boer' versus 'buur'.
1. **Hoogte.** De tong kan hoog (zoals bij /i/ in 'fiets') of laag in de mond zijn (zoals bij /a/ in 'laars').

Het Nederlands heeft 16 verschillende klinkers, die elk hun eigen combinatie van kenmerken hebben. In dit proefschrift heb ik mij gericht op twee soorten klinkercontrasten: het contrast in lipronding (ronde versus niet-ronde lippen) en het contrast in plaats (voor- versus achterklinkers).

Nagenoeg alle talen in de wereld kennen een contrast tussen voor- en achterklinkers. Meestal zijn alle achterklinkers in een taal ook rond, zoals de /o/ in 'boot', en zijn alle voorklinkers niet rond, zoals de /e/ in 'been'. De positie van de lippen en van de tong gaan dan consequent samen. Bijzonder aan het Nederlands is dat er los van een contrast tussen voor- en achterklinkers ook een contrast tussen ronde en niet-ronde klinkers is. Naast de meer gebruikelijke niet-ronde voorklinkers zoals de /e/ heeft het Nederlands namelijk ook nog ronde voorklinkers, zoals /ø/ in het woord 'leuk'. Dit soort klinkers is zeldzaam in de talen van de wereld. Door de aanwezigheid van ronde voorklinkers heeft het Nederlands een driewegcontrast tussen niet-ronde voorklinker /e/ ('been'), ronde voorklinker /ø/ ('neus') en ronde achterklinker /o/ ('boot'). Alle drie deze klinkers hebben dezelfde hoogte.

In dit onderzoek heb ik mij gericht op dit driewegcontrast tussen de Nederlandse /e/, /ø/ en /o/. Hoe zijn deze klanken gerepresenteerd in ons mentale lexicon? Hoe kunnen wij de klanken goed herkennen en uit elkaar houden? Ik heb dus onderzocht hoe het plaatscontrast (*voor* versus *achter*) en het labiaalcontrast (*rond* versus *niet-rond*) zijn gerepresenteerd bij Nederlandse klinkers. Hiertoe heb ik drie perceptie-experimenten uitgevoerd. Bij alle drie de experimenten staat de perceptie van het driewegcontrast tussen /e/, /ø/ en /o/ centraal.

In de fonologie worden de kenmerken *voor* en *achter* respectievelijk CORONAAL en DORSAAL genoemd. Ik hanteer vanaf nu deze termen. Ronde en niet-ronde klinkers worden vanaf nu respectievelijk LABIALE en niet-labiale klinkers genoemd. Met een notatie met kleine kapitalen wordt aangegeven dat het om fonologische kenmerken gaat. In Tabel 1 wordt een overzicht gegeven van de plaats- en lipkenmerken van de drie klinkers die centraal staan in dit onderzoek.

**Tabel 1.** Kenmerken van de Nederlandse klinkers /e/-/ø/-/o/

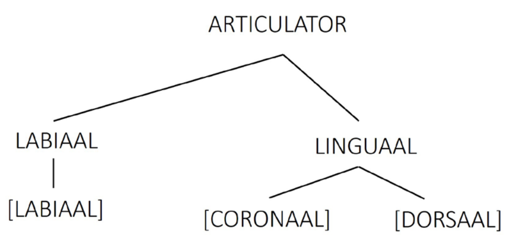
| | /e/ | /ø/ | /o/ |
|---|---|---|---|
| [CORONAL] | √ | √ | |
| [DORSAL] | | | √ |
| [LABIAL] | | √ | √ |

## Hypothese

De verwachtingen voor de perceptie van klinkercontrasten die ik heb onderzocht, zijn gebaseerd op het Featurally Underspecified Lexicon model (FUL-model), in combinatie met mijn eigen aannames voor een kenmerkgeometrie. In een kenmerkgeometrie wordt een bepaalde structuur aangenomen van fonologische kenmerken.

Het FUL-model werkt met het principe dat kenmerken die niet samen kunnen voorkomen hun knoop delen in een boomdiagram (kenmerkgeometrie). [CORONAAL] en [DORSAAL] zijn bijvoorbeeld contrastief en kunnen niet samen voorkomen. Een klinker is immers niet tegelijk een voorklinker én een achterklinker. Deze kenmerken horen dus samen onder één knoop te hangen.

In het Nederlands komen ronde achterklinkers ([LABIAAL, DORSAAL]), niet-ronde voorklinkers ([CORONAAL]) en ronde voorklinkers ([LABIAAL, CORONAAL]) voor. Het kenmerk [LABIAAL] mag dus samen voorkomen met de plaatskenmerken. Ze sluiten elkaar niet uit. Daarom heb ik gekozen voor een geometrie waarbij plaats- en lipkenmerken niet samen onder dezelfde knoop zijn geplaatst. In de geometrie in Figuur 1 neem ik twee aparte knopen aan om de labiale kenmerken en de plaatskenmerken van elkaar gescheiden te houden: de LABIALE knoop en LINGUALE knoop. Dit is anders dan de geometrie die in het FUL-model wordt aangenomen, waarbij de kenmerken [LABIAAL], [CORONAAL] en [DORSAAL] alle drie samen direct onder één knoop worden geplaatst, namelijk direct onder de ARTICULATOR knoop.

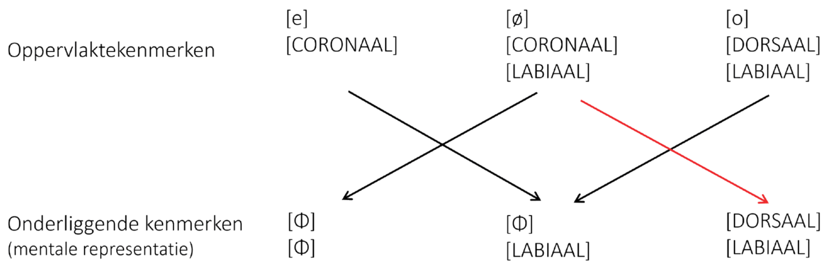



**Figuur 1.** Veronderstelde kenmerkgeometrie voor plaats- en lipkenmerken. Een aparte LABIALE knoop en LINGUALE knoop worden aangenomen, waarmee lipkenmerken en plaatskenmerken van elkaar gescheiden worden.

Merk op dat de geometrie in Figuur 1 niet een kenmerk *niet-labiaal* weergeeft. Het contrast tussen LABIAAL en niet-labiaal is in feite een contrast tussen aanwezigheid van het kenmerk LABIAAL en afwezigheid van dit kenmerk. Niet-labiaal brengt geen duidelijke akoestische kenmerken met zich mee, maar mist kenmerken die hem LABIAAL maken. Niet-labiaal wordt niet gezien als een echt fonologisch kenmerk. Dit is anders dan bij CORONAAL en DORSAAL: beide hebben typische akoestische kenmerken en worden wel als echte fonologische kenmerken gezien.

Het FUL-model neemt aan dat niet alle fonologische kenmerken worden opgeslagen in onze mentale representaties. FUL neemt CORONAAL aan als de standaard plaats van articulatie. Het FUL-model neemt daarom aan dat in alle talen van de wereld het kenmerk CORONAAL ondergespecificeerd is. Dit betekent dat in de mentale representatie dit plaatskenmerk niet gespecificeerd (ofwel opgeslagen) is, al is het wel een duidelijk aanwezig kenmerk in de akoestiek. Dit heeft tot gevolg dat de perceptie van contrasten asymmetrisch kan zijn. Dit houdt in dat een verandering van het ene naar het andere kenmerk perceptueel groter kan zijn dan een omgekeerde verandering.

Zulke asymmetrieën zijn inderdaad aangetoond in de literatuur. Bijvoorbeeld in het Duits is er een groter perceptueel verschil wanneer de dorsale klinker /o/ verandert in coronale klinker [ø] dan andersom, terwijl objectief gezien het verschil tussen twee klanken natuurlijk even groot is, ongeacht de richting van de verandering. Wanneer een CORONAAL kenmerk wordt gehoord, dan vormt dit een mismatch met een DORSAAL kenmerk dat opgeslagen is in de mentale representatie. Daarom valt een verandering van een dorsale in een coronale klank de luisteraar op. Wanneer echter een coronale klank verandert in een dorsale klank, dan worden de dorsale kenmerken wel geregistreerd, maar omdat de coronale kenmerken niet in de mentale representatie zijn opgeslagen, is daar ook geen kenmerk om mee te mismatchen. Deze verandering zorgt daarom niet voor een fonologisch conflict. Dit is samengevat in Figuur 2.




**Figuur 2.** Mapping (= vergelijking) van oppervlaktekenmerken en onderliggende kenmerken (mentale representatie) van /e/, /ø/ en /o/, weergegeven met pijlen. Zwarte pijlen geven een vergelijking aan die niet in een mismatch of conflict resulteert. De rode pijl geeft een vergelijking weer die wel tot een mismatch leidt. [Φ] geeft een ondergespecifieerd kenmerk aan: CORONAAL is afwezing in de onderliggende representatie, maar aanwezig in de oppervlaktevorm.

Om erachter te komen welke kenmerken wel en niet zijn opgeslagen in de mentale representaties, kun je dus kijken of verschillen tussen klanken asymmetrisch worden waargenomen. Dit geeft informatie over de inhoud van de representaties. Dat is wat ik bij de drie in dit proefschrift beschreven experimenten heb gedaan.

**Voorspellingen**

Voor het **plaatscontrast** (CORONAAL vs. DORSAAL, dus /ø/ vs. /o/) verwachtte ik dat CORONAAL ondergespecificeerd is en dat DORSAAL wel gespecificeerd (ofwel opgeslagen) is in de onderliggende representatie. Dit voorspelt een perceptuele asymmetrie waarbij een verandering van DORSAAL naar CORONAAL perceptueel groter is dan het omgekeerde, zoals in Figuur 2 is weergegeven.

Voor het **labiaalcontrast** (LABIAAL vs. niet-labiaal, dus /ø/ vs. /e/) voorspelde ik dat een LABIAAL oppervlaktekenmerk niet mismatcht met een niet-labiale onderliggende representatie, aangezien niet-labiaal geen echt fonologisch kenmerk is en daarom ook niet kan worden opgeslagen. Er is dus niets aanwezig in de representatie waar het kenmerk LABIAAL mee kan mismatchen. Andersom verwachtte ik dat een niet-labiale klank niet tot mapping leidt: er wordt geen kenmerk uit het signaal gehaald en er is dus niets te vergelijken met een representatie. Daarom is er in deze richting ook geen mismatch. Perceptie zou dan symmetrisch zijn. Dit zou in overeenstemming zijn met resultaten die gevonden zijn bij Duitstalige luisteraars. Het is echter raadselachtig hoe Nederlandstalige luisteraars dan in staat zouden zijn op efficiënte wijze dit contrast waar te nemen dat in het Nederlands toch erg belangrijk is. Ik was dus benieuwd of deze voorspelling ook daadwerkelijk uit zou komen.

Voor het **gecombineerde contrast** (CORONAAL niet-labiaal vs. DORSAAL LABIAAL, dus /e/ versus /o/) voorspelde ik een asymmetrie op basis van de plaatskenmerken. Wederom zou CORONAAL ondergespecificeerd zijn en DORSAAL wel opgeslagen zijn, waardoor een verandering van dorsaal in coronaal meer zou opvallen dan andersom. Er werd geen belangrijke rol van het labiaalcontrast voorspeld voor de uitkomsten van het gecombineerde contrast, omdat voor het labiaalcontrast symmetrie werd aangenomen.

**Experimenten**

In **Hoofdstuk 2** beschrijf ik mijn onderzoek naar de perceptie van klinkerveranderingen bij Nederlandstalige volwassenen met behulp van elektro-encefalografie (EEG). Met EEG wordt de elektrische hersenactiviteit geregistreerd. Met EEG heb ik de reactie van het brein op klinkerveranderingen gemeten: de *mismatch negativity* respons (MMN). De MMN-respons is een automatische reactie van het brein op een klankverandering. Deze respons geeft inzicht in de perceptuele verschillen tussen spraakklanken: een perceptief groot verschil zorgt voor een grotere en/of snellere reactie van het brein dan een perceptief klein verschil.

Tijdens het experiment hoorden de deelnemers een stroom van losse klanken. Zij hoorden bijvoorbeeld telkens opnieuw een /o/, maar deze stroom werd soms onderbroken door een andere klank, bijvoorbeeld een [e]. De luisteraar verwacht dan een /o/, maar hoort een [e]. De kenmerken van de afwijkende klinker (kenmerken uit de akoestiek) worden vergeleken met de kenmerken van de verwachte klinker (kenmerken in de representatie). Wanneer wij de respons vergelijken tussen een verandering van bijvoorbeeld een dorsale naar coronale klank en andersom, kunnen wij kijken of de reactie in beide richtingen verschilt en dus asymmetrisch is of niet. Dit geeft informatie over of een kenmerk aanwezig is in de mentale representatie.

Wat betreft het plaatscontrast tussen dorsale /o/ en coronale /ø/ lieten de uitkomsten inderdaad een asymmetrie zien. Er was een grotere MMN respons wanneer een dorsale klank in een coronale klank veranderde dan andersom. Dit was een verwachte uitkomst, want de asymmetrie was in overeenstemming met de aanname van FUL dat [CORONAAL] ondergespecificeerd is en [DORSAAL] wel wordt opgeslagen in de representatie.

Voor het labiaalcontrast (LABIAAL versus niet-labiaal, dus /ø/ versus /e/) werd daarentegen geen asymmetrie verwacht. Voor beide richtingen van het contrast werd verwacht dat er geen mismatch zou zijn. Wanneer een [LABIAAL] kenmerk uit de akoestiek van [ø] gehaald wordt, kan deze niet mismatchen met een niet-labiale representatie van /e/, omdat dit niet is opgeslagen in de representatie. Het is immers geen echt fonologisch kenmerk. In tegengestelde richting werd ook geen mismatch voorspeld. In dat geval hoort de luisteraar namelijk een niet-labiale klank, maar afwezigheid van een kenmerk is niet iets dat uit het signaal gehaald kan worden en daarom lokt het horen van een niet-labiale klank niet uit dat dit kenmerk wordt vergeleken met een onderliggend kenmerk. Er is immers geen kenmerk uit het signaal gehaald en er valt dus niets te vergelijken. Er is dus ook geen mismatch. Tegen de verwachting in werd er wél een asymmetrie gevonden. Er was een grotere MMN respons wanneer een niet-labiale klank in een labiale klank veranderde dan andersom.

Om na te gaan of de bevindingen uit het EEG-experiment repliceerbaar zijn, heb ik dezelfde contrasten tussen /e/, /ø/ en /o/ bij Nederlandstalige volwassenen ook onderzocht met een andere methode. In **Hoofdstuk 3** heb ik een semantisch priming experiment beschreven. In dit experiment hoorden deelnemers een woord en daarna lazen zij een woord. Zij hadden de taak zo snel mogelijk te beslissen of het woord dat zij lazen een bestaand Nederlands woord was of niet. Reactietijden worden beïnvloed door het woord dat zij hoorden. Hoor je eerst *boot* en lees je dan *schip* dan ben je sneller met reageren dan wanneer je eerst *bril* hoort en dan *schip.* In het eerste geval is het woord dat je hoorde namelijk min of meer synoniem aan het woord dat je las. Vanwege deze relatie in betekenis was het woord dat je las eigenlijk in je hoofd al actief geworden door het woord dat je hoorde. Daardoor kun je het gelezen woord sneller herkennen dan wanneer het woord dat je hoort niets te maken heeft met het woord dat je leest. Zulke versnellingen van reactietijden noemen we priming-effecten.

In het experiment werd de uitspraak van de woorden die de deelnemers hoorden gemanipuleerd: soms werd het woord met een verkeerde klinker uitgesproken, bijv. **beut* in plaats van *boot*. Wanneer deze verkeerde klinker niet zorgt voor een mismatch, dan is de aanname dat je nog steeds toegang krijgt tot het woord in je mentale lexicon. Dit betekent dat je het 'bedoelde' woord nog steeds herkent. Dan wordt het woord *boot* bijvoorbeeld nog steeds geactiveerd. Dan verwacht je nog steeds een snellere reactie op een gelezen synoniem. Wanneer de verkeerde klinker echter zorgt voor een mismatch met de kenmerken van het 'bedoelde' woord, dan is er geen woordherkenning. Daarmee is er ook geen toegang tot de betekenis van het woord en worden ook synoniemen niet alvast geactiveerd. Er is dan geen semantisch priming effect (= versnelling door betekenisrelatie). Door reactietijden te bestuderen kun je met deze methode dus ook onderzoeken of er asymmetrieën zijn in de klinkercontrasten.

Helaas was er na het opschonen van de data voor de analyse niet genoeg over om statistisch significante effecten te kunnen ontdekken. Wat betreft het plaatscontrast geven de uitkomsten een vergelijkbaar asymmetrisch patroon als het EEG-experiment van Hoofdstuk 2. De resultaten zijn echter onduidelijk wat betreft het labiaalcontrast.



Naast het bestuderen van perceptie bij volwassenen kan ook kennis over hoe het volwassen systeem wordt opgebouwd door jonge kinderen die hun moedertaal verwerven inzicht geven in het volwassen systeem. In **Hoofdstuk 4** beschrijf ik het onderzoek naar de perceptie van de contrasten tussen /e/, /ø/ en /o/ bij 151 driejarigen uit de regio Nijmegen die het Nederlands als moedertaal verwerven. Op de leeftijd van drie jaar beginnen kinderen ronde voorklinkers te produceren. Het is daarom een interessante leeftijd om te kijken naar hoe hun representaties er in dit opzicht uitzien.

Tijdens het experiment leerden kinderen eerst een nieuw woord voor een object dat zij nog niet kenden (zie Hoofdstuk 4 voor afbeeldingen van de objecten). Daarna werd getest of zij dit nieuwe woord herkenden. Zij keken naar een televisiescherm waarop twee objecten te zien waren. Een daarvan werd benoemd met het woord dat zij net hadden geleerd. Er werd gemeten of kinderen het woord herkenden door te meten of het kind naar het juiste plaatje keek. Soms werd het woord verkeerd uitgesproken: dan werd het met een andere klinker uitgesproken dan hoe het kind het geleerd had. De aanname is net als bij het semantisch priming experiment in Hoofdstuk 3 dat wanneer deze verkeerde uitspraak niet zorgt voor een mismatch, het oorspronkelijke woord nog steeds te herkennen is. Wanneer een klankverandering wel zorgt voor een mismatch met de representatie van het geleerde woord, dan wordt het woord niet meer herkend. Of het woord wordt herkend, valt af te leiden uit het kijkgedrag van de kinderen.

Hierin zijn dus asymmetrieën te onderzoeken: maakt het uit in welke richting een klank verandert voor de woordherkenning van de kinderen?

De resultaten toonden aan dat de kinderen de nieuwe woorden succesvol hadden geleerd. Daarnaast waren de kinderen gevoelig voor alle contrasten die ik heb getest. Wanneer woorden verkeerd werden uitgesproken, viel dit verschil de kinderen op. Het maakte daarbij niet uit om welk contrast het ging. Kinderen van drie jaar oud nemen het Nederlandse driewegcontrast tussen achterklinkers, voorklinkers en ronde voorklinkers dus goed waar.

Opvallend was dat er geen asymmetrieën gevonden zijn bij de peuters. Het maakte dus niet uit in welke richting een fonologisch kenmerk veranderde. Ongeacht de richting van verandering zorgde een verkeerd uitgesproken woord ervoor dat er geen woordherkenning meer plaatsvond. Waarschijnlijk was de gebruikte methode niet gevoelig genoeg om asymmetrieën te kunnen meten, omdat de kinderen al zeer vaardig waren in het waarnemen van de verschillen. Om zeker te weten of de perceptie van kinderen daadwerkelijk symmetrisch is, zou een gevoeligere methode gebruikt moeten worden, zoals EEG, waarbij direct de reactie van het brein gemeten wordt. De methode die nu gebruikt is, heeft wel bij leeftijdgenoten goed gewerkt, maar dat was bij andere contrasten. De door mij gebruikte contrasten waren misschien te eenvoudig voor deze leeftijdsgroep.

**Interpretatie**

De uitkomsten van het EEG-experiment in Hoofdstuk 2 zijn het meest betekenisvol. Daarom is de interpretatie met name op deze uitkomsten gebaseerd.

Er bleek een plaatsasymmetrie te zijn, zoals ook voorspeld was op basis van onderspecificatie van het kenmerk [CORONAAL] en specificatie van het kenmerk [DORSAAL]. De perceptuele asymmetrie tussen /ø/ en /o/ is in overeenstemming met wat ook in eerdere literatuur gevonden is voor bijvoorbeeld deze klinkers in het Duits en het Frans. Deze uitkomst was dus niet verrassend.

Daarnaast vond ik een asymmetrie tussen labiale klinker /ø/ en niet-labiale klinker /e/. Deze asymmetrie was niet voorspeld. In het FUL-model gaat men ervanuit dat wanneer een bepaald kenmerk niet aanwezig is in de onderliggende representatie, daar ook niets mee kan mismatchen. De /e/ heeft geen lipronding, dus zou een LABIAAL kenmerk van [ø] nergens mee kunnen mismatchen en er zou in dit geval dus geen fonologisch conflict moeten zijn. De gevonden asymmetrie betekent echter dat er op fonologisch niveau toch een conflict is, wanneer een niet-labiale klinker /e/ verandert in labiale klinker [ø]. Behalve de aangenomen kenmerkgeometrie die lipkenmerken apart houdt van plaatskenmerken (zie Figuur 1) is er dus meer nodig om te verklaren hoe de klinkers verwerkt worden. Ik doe in Hoofdstuk 5 een voorstel voor hoe dit zou kunnen werken.

Ik stel voor dat er bepaalde regels gebruikt worden tijdens het vergelijken van wat je hoort met wat je hebt opgeslagen in je hoofd: zogenaamde *'mapping rules'*. Deze regels definiëren welke kenmerken uit de spraak die je hoort moeten worden vergeleken met welke kenmerken in de representatie wanneer je naar spraak luistert. Zo is het mappen efficiënt en betekenisvol: er worden geen kenmerken met elkaar vergeleken als dat niet zinnig is – zoals indien de kenmerken gecombineerd kunnen worden, zoals [CORONAAL] en [HOOG] (bijvoorbeeld bij /i/ in fiets).

Ik stel in dit proefschrift voor dat deze *mapping rules* die vertellen wat er met wat vergeleken moet worden, gebruik maken van de kenmerkgeometrie: de structuur van de verschillende kenmerken. De knopen in de structuur geven immers aan welke kenmerken contrastief zijn, zoals in het geval van [CORONAAL] en [DORSAAL] duidelijk is: deze kenmerken zijn contrastief en delen de bovenliggende LINGUALE knoop. De aanname dat knopen een rol spelen leek impliciet al aanwezig in het FUL-model, waar men er ook van uitgaat dat een kenmerk als [CORONAAL] niet vergeleken wordt met een kenmerk als [HOOG], waar het niet mee contrasteert.

6

Het FUL-model neemt aan dat knopen in de geometrie alleen aanwezig zijn in de representatie, wanneer er een fonologisch dochterkenmerk aanwezig is dat daarbij hoort. De consequenties hiervan verschillen voor het labiaalcontrast enerzijds en het plaatscontrast anderzijds. Bij het labiaalcontrast is er een contrast tussen een fonologisch kenmerk ([LABIAAL]) en afwezigheid van het kenmerk (niet-labiaal). Een niet-labiale klank heeft dus géén LABIALE knoop. Alleen een labiale klank heeft ook een LABIALE knoop in diens geometrie. Bij het plaatscontrast is er een contrast tussen twee verschillende fonologische kenmerken: [CORONAAL] en [DORSAAL]. De bovenliggende LINGUALE knoop is er, ongeacht welke van beide kenmerken aanwezig is. Dit is een cruciaal verschil tussen beide contrasten.

Voor het plaatscontrast kan mapping van kenmerken plaatsvinden met behulp van *mapping rules*, waarbij bijvoorbeeld een [CORONAAL] kenmerk wordt vergeleken met een [DORSAAL] kenmerk, omdat de *mapping rule* zegt dat het binnengekomen LINGUALE kenmerk vergeleken moet worden met het opgeslagen LINGUALE kenmerk. De aanname dat zulke mapping rules een rol spelen, verandert de voorspellingen voor de perceptie van het plaatscontrast niet en is dus in lijn met de uitkomsten van het onderhavige onderzoek.

Voor het labiaalcontrast heeft de aanname van mapping rules met een actieve rol voor de knopen in de geometrie weldegelijk consequenties voor de perceptie. Wanneer een labiale klank wordt gehoord zoals [ø], dan heeft die klank een LABIALE knoop en wil je dus het binnengekomen kenmerk vergelijken met het kenmerk dat in de onderliggende representatie ook onder de LABIALE knoop zit. Wanneer je dan een labiale klank [ø] vergelijkt met de onderliggende representatie van een niet-labiale klank /e/, loop je tegen een probleem aan. De /e/ heeft namelijk geen LABIALE knoop. Het lukt dan dus niet om een kenmerk te vinden dat bij die knoop hoort en er kunnen dus geen fonologische kenmerken met elkaar vergeleken worden. Dit betekent dat er eigenlijk al een conflict is tussen de geometrie van de gehoorde klank [ø] en die van de opgeslagen klank /e/. Wanneer een /e/ veranderd is in een [ø], probeer je het LABIALE kenmerk van [ø] dus te mappen met iets, maar dat blijkt onmogelijk, omdat de juiste plek om te mappen er niet is: de knoop is er niet. Dit resulteert in een groot perceptueel verschil tussen /e/ en /ø/ in deze richting van het contrast.

*Mapping rules* die gebruik maken van de knopen in de geometrie kunnen een verklaring bieden voor de gevonden labiaal-asymmetrie. Zoals hierboven beschreven, is een fonologisch conflict te verwachten wanneer een niet-labiale klank verandert in een labiale klank. In tegenovergestelde richting is er geen fonologisch conflict te verwachten. Wanneer een niet-labiale klank zoals /e/ gehoord wordt, wordt uit dit signaal niets gehaald dat gemapt hoeft te worden op de LABIALE knoop. De /e/ heeft immers geen LABIAAL kenmerk en ook geen LABIALE knoop. Er wordt dus geen mapping van labiale kenmerken geïnitieerd. Dan kan daar ook geen conflict uitkomen.

Samenvattend stel ik voor dat mapping rules definiëren welke oppervlaktekenmerken op welke onderliggende kenmerken worden gemapt. Deze mapping maakt gebruik van de knoopstructuur in de geometrie. De geometrie kan in combinatie met mapping rules een verklaring bieden voor zowel de plaatsasymmetrie als de labiaalasymmetrie, gevonden voor Nederlandse klinkers. Het biedt een verklaring voor hoe verschillende soorten contrasten kunnen worden waargenomen. Het plaatscontrast resulteert in een mismatch tussen twee kenmerken en daarom in een fonologisch conflict. Wanneer er echter een contrast is tussen een kenmerk en afwezigheid van een kenmerk, zoals bij het labiaalcontrast, dan kan een fonologisch conflict optreden op het niveau van de geometrie, nog voordat twee kenmerken vergeleken kunnen worden.

De labiaalasymmetrie is in dit onderzoek gevonden bij Nederlandse luisteraars, maar deze asymmetrie was niet aanwezig bij Duitse luisteraars in een eerder onderzoek. De klankinventaris van beide talen lijkt op elkaar, maar toch zijn beide taalsystemen verschillend. Ronde voorklinkers lijken een andere rol te spelen in het Duits dan in het Nederlands (zie hoofdstuk 5). Verschillen tussen het Duitse en het Nederlandse klanksysteem kunnen zich manifesteren in de mapping rules die beide talen gebruiken. Het labiaalcontrast lijkt in het Duits minder belangrijk om op fonologisch niveau te maken. Het kan zijn dat er in het Duits geen mapping rule aanwezig is die zegt dat een labiaal oppervlaktekenmerk gemapt moet worden. Er wordt dan misschien geen mapping geïnitieerd. Dan zou je verwachten dat er in het Duits symmetrische waarneming is van /ø/ versus /e/ en dat blijkt inderdaad het geval te zijn. Als verschillende talen verschillende mapping rules kunnen hebben, kunnen verschillende talen ook met dezelfde kenmerkgeometrie en dezelfde fonologische kenmerken andere perceptuele patronen laten zien.

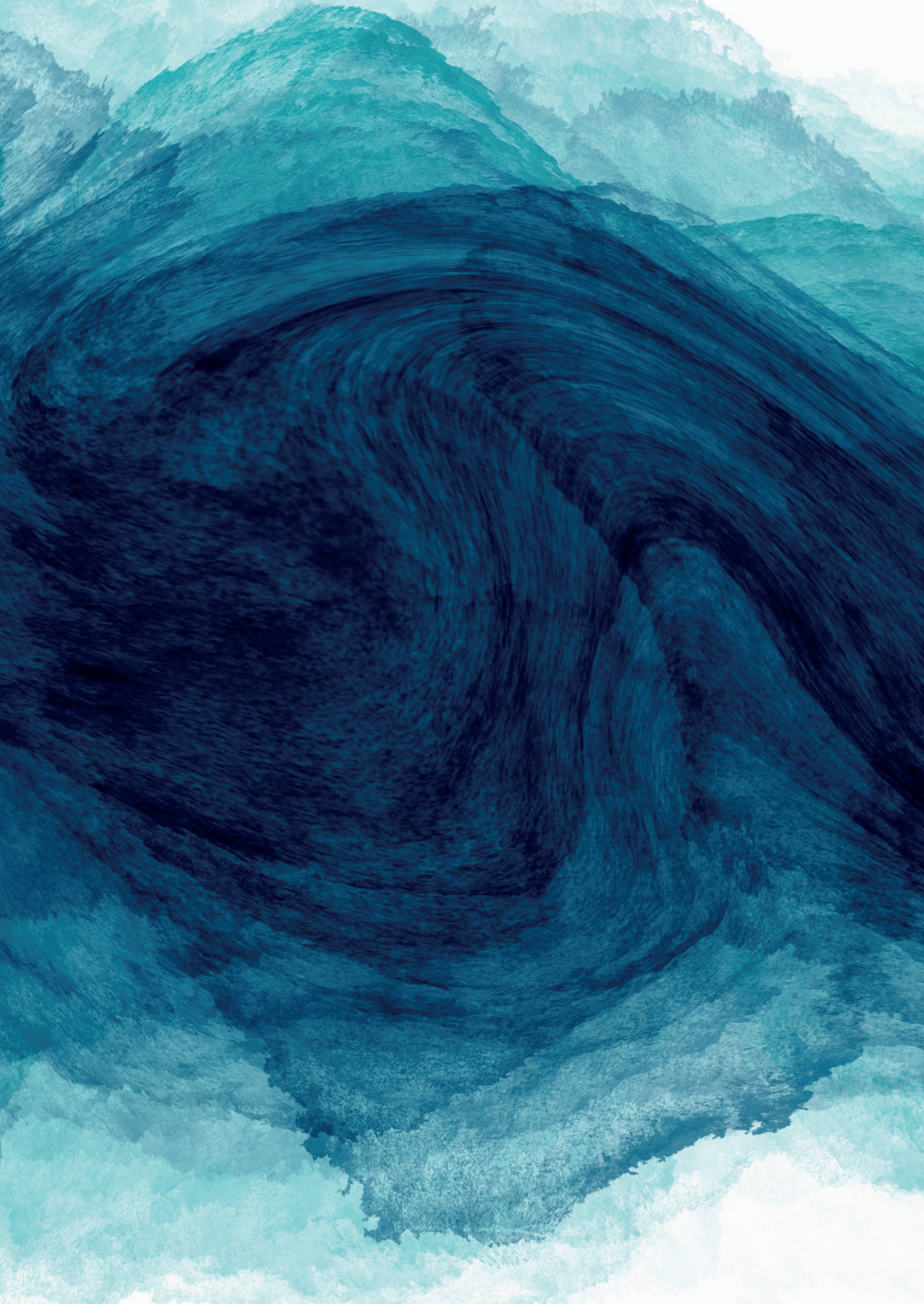
**Conclusie**

De resultaten van de experimenten in dit proefschrift tonen asymmetrieën aan in de klinkerperceptie in het Nederlands bij zowel het plaatscontrast als het labiaalcontrast. De aanwezigheid van asymmetrieën is in lijn met de aanname dat niet alle mogelijke kenmerken daadwerkelijk worden opgeslagen in de mentale representaties, maar dat sommige kenmerken ondergespecificeerd kunnen zijn.

Daarnaast tonen de resultaten dat een fonologisch conflict op meerdere manieren kan ontstaan. Naast een fonologisch conflict vanwege een *mapping mismatch* (zoals bij het plaatscontrast), kan er ook een fonologisch conflict optreden wanneer mapping van bepaalde kenmerken überhaupt niet mogelijk blijkt, wanneer de knoopstructuur in de geometrie afwijkt (zoals bij het labiaalcontrast).

Het lijkt er dus op dat er verschillende typen fonologisch contrast te onderscheiden zijn. Vervolgonderzoek zou moeten uitwijzen of dit ook gevonden kan worden bij andere talen dan het Nederlands. Daarnaast zou vervolgonderzoek kunnen uitwijzen of bij vergelijkbare contrasten waarbij er ook een contrast is tussen een kenmerk en afwezigheid van een kenmerk (zoals *stemhebbend* versus *stemloos*) vergelijkbare patronen optreden als in het huidige experiment het geval is voor het labiaalcontrast. Voor het in kaart brengen van de totstandkoming van het volwassen Nederlandse systeem, inclusief geometrie en mapping rules, zou het waardevol zijn een longitudinaal onderzoek uit te voeren dat de taalverwerving op dit gebied volgt.

Het Nederlands en het Duits lijken in eerste instantie op elkaar wanneer je kijkt naar welke klinkers in beide talen voorkomen, maar in dit onderzoek is naar voren gekomen dat vergelijkbare klinkercontrasten in beide talen toch verschillend worden waargenomen. Ook wanneer fonologische contrasten aan de oppervlakte vergelijkbaar lijken, kunnen de onderliggende taalsystemen verschillen. Gedetailleerde perceptie-experimenten die verschillende talen vergelijken zouden meer inzicht kunnen bieden in de overeenkomsten en verschillen tussen verschillende talen en taalsystemen. Op die manier kan ook meer inzicht verkregen worden in of bepaalde fonologische verschijnselen algemeen of taalspecifiek zijn. Oog voor detail is dus belangrijk voor inzicht in het grote geheel.



# Bibliography

Adank, P., van Hout, R., & Smits, R. (2004). An acoustic description of the vowels of Northern and Southern Standard Dutch. *The Journal of the Acoustical Society of America*, *116*(3), 1729-1738. https://doi.org/10.1121/1.1779271

Altmann, G.T.M., & Kamide, J. (2004). Now You See It, Now You Don't: Mediating the Mapping between Language and the Visual World. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 347–386). Psychology Press, New York.

Altvater-Mackensen, N., & Mani, N. (2013). Word-form familiarity bootstraps infant speech segmentation. *Developmental Science*, *16*(6), 980-990. https://doi.org/10.1111/desc.12071

Altvater-Mackensen, N., van der Feest, S.V.H., & Fikkert, P. (2014). Asymmetries in early word recognition: The case of stops and fricatives. *Language Learning and Development*, *10*(2), 149–178. https://doi.org/10.1080/15475441.2013.808954

Anderson, J.L., Morgan, J.L., & White, K.S. (2003). A statistical basis for speech sound discrimination. *Language and Speech, 46*(2), 155-182. doi:10.1177/00238309030460020601

Archangeli, D. (1988). Aspects of underspecification theory. *Phonology*, *5*(2), 183–207. doi:10.1017/S0952675700002268

Baayen, R.H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (CD-ROM). Linguistic Data Consortium, Philadelphia: University of Pennsylvania.

Ballem, K., & Plunkett, K. (2005). Phonological specificity in children at 1;2. *Journal of child language*, *32*. 159-73. doi:10.1017/S0305000904006567.

Bailey, T.M., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development*, *17*(2), 1265-1282. https://doi.org/10.1016/S0885-2014(02)00116-8

Beers, M. (1995). The Phonology of Normally Developing and Language-impaired Children. PhD Thesis. University of Amsterdam, Institute for Functional Research into Language and Language Use (IFOTT).

Bergelson, E., & Swingley, D. (2012). At 6–9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences*, *109* (9), 3253-3258; DOI: 10.1073/pnas.1113380109

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 speech research* (pp. 171–204). Timonium, MD: York Press.

Best, C.T. (1994). The emergence of native-language phonological influences in infants: a perceptual assimilation model. In J. Goodman & H. Nusbaum (Eds.), *The development of speech perception: the transition from speech to spoken words* (pp. 167-224). Cambridge, MA: MIT Press.

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

Best, C.T., Goldstein, L.M., Nam, H., & Tyler, M.D. (2016). Articulating What Infants Attune to in Native Speech. *Ecological Psychology*, *28*(4), 216–261. https://doi.org/10.1080/10407413.2016.1230372

Best, C.C., & McRoberts, G.W. (2003). Infant Perception of Non-Native Consonant Contrasts that Adults Assimilate in Different Ways. *Language and Speech*, *46*(2–3), 183–216. https://doi.org/10.1177/00238309030460020701

Boersma, P., & Weenink, D. (2012). PRAAT: doing phonetics by computer [Computer program]. Version 5.3.22, http://www.praat.org/

Bölte, J., & Coenen, E. (2000). Domato primes paprika: Mismatching pseudowords activate semantic and phonological representations, *SWAP-2000*, 59-62.

Bohn, O.-S., & Polka, L. (2001). Target spectral, dynamic spectral, and duration cues in infant perception of German vowels. *The Journal of the Acoustical Society of America*, *110*(1), 504–515. https://doi.org/10.1121/1.1380415

Booij, G. (1981). Generatieve fonologie van het Nederlands. Utrecht: Het Spectrum.

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

Browman, C.P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, 6, 201–251.

Brown, C. (1997). Acquisition of segmental structure: consequences for speech perception and Second Language Acquisition. Doctoral dissertation, McGill University, Montréal, Québec, Canada.

Brown, C., & Matthews, J. (1997). The role of feature geometry in the development of phonemic contrasts. In S.J. Hannahs, & M. Young-Scholten (Eds), *Focus on Phonological Acquisition.* John Benjamins Publishing Company Amsterdam/Philadelphia.

Buckler, H., & Fikkert, P. (2016). Dutch and German 3-year-old's representations of voicing alternations. *Language and Speech*, *59*(2), 236- 265.

Bybee, J. (2001). *Phonology and language use*. Cambridge: Cambridge University Press.

Castor EDC. (2015). Castor Electronic Data Capture. [online]. Available at: https://castoredc.com.

Charles-Luce, J., & Luce, P.A. (1990). Similarity neighbourhoods of words in young children's lexicons. *Journal of Child Language,* 17(1), 205–215. https://doi.org/10.1017/S0305000900013180

Charles-Luce J., & Luce P.A. (1995) An examination of similarity neighbourhoods in young children's receptive vocabularies. *Journal of Child Language,* 22(3). 727-35. doi: 10.1017/s0305000900010023. PMID: 8789521.

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

Goldinger, S. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review*, 105, 251–79. https://doi:10.1037/0033-295X.105.2.251

Clements, G.N. (1976). Palatalization: linking or assimilation? *CLS.* 12. 96-106.

Clements, G.N. (1985). The geometry of phonological features. *Phonology Yearbook*, *2*(1), 225–252. https://doi.org/10.1017/S0952675700000440

Clements, G.N. (1999). The geometry of phonological features. In J. Goldsmith, (Ed.), *Phonological theory: the essential readings* (pp. 201–223). Oxford: Blackwell..

Clements, G.N., & Hume, E. (1995). The internal organization of speech sounds. In J. Goldsmith (Ed.), *Handbook of phonological theory* (pp. 245-306). Oxford: Basil Blackwell.

Coady, J., & Aslin, R. (2003). Phonological neighbourhoods in the developing lexicon. *Journal of Child Language*, *30*. 441-69.

Cornell, S.A., Lahiri, A., & Eulitz, C. (2011). "What you encode is not necessarily what you store": Evidence for sparse feature representations from mismatch negativity. *Brain Research, 1394,* 79–89. doi:10.1016/j.brainres.2011.04.001

Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of working memory: mechanisms of active maintenance and executive control.* Cambridge: Cambridge University Press.

Curtin, S., Fennell, C., & Escudero, P. (2009). Weighting of vowel cues explains patterns of word object associative learning. *Developmental Science*, *12*(5), 725–731. https://doi.org/10.1111/j.1467-7687.2009.00814.x

Dahan, D., Magnuson, J.S., Tanenhaus, M., & Hogan, E.M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes, 16*, 507-534.

Davis, B.L. & MacNeilage, P.F. (1990). Acquisition of correct vowel production: A quantitative case study. *Journal of Speech and Hearing Research*, *33*, 16–27.

Davis, B.L., & MacNeilage, P.F. (1995). The articulatory basis of babbling. *Journal of Speech & Hearing Research, 38*(6), 1199–1211. https://doi.org/10.1044/jshr.3806.1199

Davis, B., MacNeilage, P., & Matyear, C. (2002). Acquisition of serial complexity in speech production: A comparison of phonetic and phonological approaches to first word production. *Phonetica*, *59*, 75-107. doi:10.1159/000066065.

De Groot, A.W. (1931). De Wetten der Phonologie en hun Betekenis voor de Studie van het Nederlands, *Nieuwe Taalgids, 25,* 225-239.

De Jonge, M.J.I., & Boersma, P. (2015). French high-mid vowels are underspecified for height. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow: The University of Glasgow.

Dehaene-Lambertz, G. (1997). Electrophysiological correlates of categorical phoneme perception in adults. *Neuroreport*, *3;8*(4). 919-24. https://doi:10.1097/00001756-199703030-00021.

Delle Luche, C., Durrant, S., Poltrock, S., & Floccia, C. (2015). A methodological investigation of the Intermodal Preferential Looking paradigm: Methods of analyses, picture selection and data rejection criteria. *Infant Behavior & Development, 40*, 151–172. https://doi.org/10.1016/j.infbeh.2015.05.005

Diesendruck., G., & Markson, L. (2001). Children's avoidance of lexical overlap: A pragmatic account. *Developmental Psychology*, *37*, 630-641.

Dijkstra, C., & Fikkert, P. (2011). Universal Constraints on the Discrimination of Place of Articulation? Asymmetries in the Discrimination of 'paan' and 'taan' by 6-month-old Dutch Infants. In: Nick Danis, Kate Mesh, and Hyunsuk Sung (Eds.), *Proceedings of the Boston University Conference on Child Development, 35*, 170-182.

Dresher, B.E. (2009). *The contrastive hierarchy in phonology*. (Cambridge Studies in Linguistics 121.) Cambridge: Cambridge University Press.

Dresher, B.E. (2015). The motivation for contrastive feature hierarchies in phonology. *Linguistic Variation*, *15*(1), 1–40. https://doi.org/10.1075/lv.15.1.01dre

Dresher, B.E. (2018). *Contrastive Feature Hierarchies in Synchronic and Diachronic Phonology*. 8.

Durrant, S. (2014). The influence of long-term exposure to dialect variation on representation specificity and word learning in toddlers. Doctoral dissertation, Plymouth University.

Eijkman, L.P.H. (1937). Phonetiek van het Nederlands. Haarlem: Bohn.

Ernestus, M. (2014). Acoustic reduction and the roles of abstractions and exemplars in speech processing. *Lingua, 142,* 27-41. https://doi.org/10.1016/j.lingua.2012.12.006

Eulitz, C., Diesch, E., Pantev, C., Hampson, S., & Elbert, T. (1995). Magnetic and electric brain activity evoked by the processing of tone and vowel stimuli. *Journal of Neuroscience*, 15, 2748–2755.

Eulitz, C., & Lahiri, A. (2004). Neurobiological Evidence for Abstract Phonological Representations in the Mental Lexicon during Speech Recognition. *Journal of Cognitive Neuroscience*, *16*(4), 577–583. https://doi.org/10.1162/089892904323057308

Feldman, N.H., Griffiths, T. L., & Morgan, J. L. (2009). The influence of categories on perception: Explaining the perceptual magnet effect as optimal statistical inference. *Psychological Review, 116*(4), 752–782. https://doi.org/10.1037/a0017196

Fennell, C., & Werker, J. (2003). Early Word Learners' Ability to Access Phonetic Detail in Well-Known Words. *Language and speech, 46*, 245-64. 10.1177/00238309030460020901.

Fernald, A., Zangl, R., Portillo, A.L., & Marchman, V.A. (2008). Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In I. A. Sekerina, E. M. Fernández, & H. Clahsen (Eds*.), Language acquisition and language disorders: Vol. 44. Developmental psycholinguistics: On-line methods in children's language processing* (pp. 97–135). Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/lald.44.06fer

Fikkert, P. (2010). Developing representations and the emergence of phonology: Evidence from perception and production. In C. Fougeron, B. Kühnert, M. d'Imperio, & N. Vallée (Eds.), *Laboratory phonology 10: Variation, phonetic detail and phonological representation* (Phonology & Phonetics 4-4) (pp. 227–258). Berlin. Germany: Mouton.

Fikkert P., & Levelt C.C. (2008). How does Place fall into place? The lexicon and emergent constraints. In: Avery P., Dresher B.E., Rice K. (red.) *Contrast in Phonology; Theory, Perception, Acquisition* (pp. 231-270). Berlin: Mouton de Gruyter.

Golinkoff, R.M., Hirsh-Pasek, K., Cauley, K.M., & Gordon, L. (1987): The eyes have it: Lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*, 23-45.

Golinkoff, R.M., Ma, W., Song, L., & Hirsh-Pasek, K. (2013). Twenty-Five Years Using the Intermodal Preferential Looking Paradigm to Study Language Acquisition: What Have We Learned? *Perspectives on Psychological Science*, *8*(3), 316–339. https://doi.org/10.1177/1745691613484936

Grieser, D., & Kuhl, P.K. (1989). Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology*, *25*(4), 577–588. https://doi.org/10.1037/0012-1649.25.4.577

Gussenhoven, C. (1992). Dutch. *Journal of the International Phonetic Association,* 22. 45 - 47. 10.1017/S002510030000459X.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition,* 87, B23-B34

Hestvik, A., & Durvasula, K. (2016). Neurobiological evidence for voicing underspecification in English. *Brain Lang*, 152, 28–43. doi:10.1016/j.bandl. 2015.10.007.

Hilton, M., & Westermann, G. (2016). The effect of shyness on children's formation and retention of novel word-object mappings. *Journal of Child Language, 44*(6), 1394-1412. doi:10.1017/S030500091600057X

Houston-Price C, Plunkett, K., & Harris P. (2005). Word-learning wizardry at 1;6. *Journal of Child Language, 32*(1), 175-89. doi: 10.1017/s0305000904006610. PMID: 15779882.

Huettig, F., & Altmann, G. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition. 96*. B23-32. 10.1016/j.cognition.2004.10.003.

Iverson, P., & Kuhl, P.K. (1996). Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. *Journal of the Acoustical Society of America*, *99*(2), 1130–1140. https://doi.org/10.1121/1.415234

Jakobson, R. (1931/62). Phonemic notes on Standard Slovak. In R. Jakobson (1962), *Selected writings I. Phonological studies* (pp. 221–230). The Hague: Mouton.

Jakobson, R. (1941/68). Child language, aphasia, and phonological universals. The Hague: Mouton. English translation of Kindersprache, aphasie und allgemeine lautgesetze. Uppsala.

Jakobson, R., Fant. C.G.M., & Halle. M. (1952). Preliminaries to speech analysis: The distinctive features and their correlates. Cambridge: M.I.T. Press.

Jakobson, R., & Morris, H. (1956). Fundamentals of Language. The Hague: Mouton

Johnson, K. (1997). Speech perception without speaker normalization: an exemplar model. In Johnson, K., Mullenix, J.W. (Eds.), *Talker variability in speech processing* (pp. 145–166). San Diego: Academic Press.

JMP®, Version 15. SAS Institute Inc., Cary, NC, 1989-2018.

Jusczyk P.W., Cutler A., & Redanz N.J. (1993) Infants' preference for the predominant stress patterns of English words. *Child Development, 64*(3), 675-87. PMID: 8339688.

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. https://doi.org/10.1006/jmla.1994.1030

Kalashnikova, M., Mattock, K., & Monaghan, P. (2015). The effects of linguistic experience on the flexible use of mutual exclusivity in word learning. *Bilingualism: Language and Cognition, 18*(4), 626-638.

Keating, P. A. (1988). Underspecification in Phonetics. *Phonology*, *5*(2), 275–292. https://doi.org/10.1017/S095267570000230X

Keuleers, E., Brysbaert, M. & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods, 42*, 643–650. https://doi.org/10.3758/BRM.42.3.643

Kern, S., & Davis, B. L. (2009). Emergent complexity in early vocal acquisition: Cross linguistic comparisons of canonical babbling. *Approaches to Phonological Complexity* (pp.353-376). Berlin, New York: De Gruyter Mouton. https://doi.org/10.1515/9783110223958.353

Kern, S., Davis, B. L. & Zink, I. (2010). From babbling to first words in four languages: Common trends, cross language and individual differences. In: J.M. Hombert & F. d'Errico (Eds.), *Becoming Eloquent* (pp. 205-232). Cambridge, UK: John Benjamins Publishers.

Keyser, S., & Stevens, K. (1994). Feature Geometry and the Vocal Tract. *Phonology, 11*(2), 207–236.

Kotzor, S. & Wetterlin, A. & Roberts, A. & Lahiri, A. (2015). Processing of Phonemic Consonant Length: Semantic and Fragment Priming Evidence from Bengali. *Language and Speech*, *59*(1), 83–112. https://doi.org/10.1177/0023830915580189

Kotzor, S., Wetterlin, A., & Lahiri, A. (2017). Symmetry or asymmetry: Evidence for underspecification in the mental lexicon. In Lahiri, A. & Kotzor, S (Eds.), *The Speech Processing Lexicon* (pp. 85-106). Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110422658-005

Kuhl, P.K. (1991). Human adults and human infants show a "Perceptual Magnet Effect" for the prototypes of speech categories: Monkeys do not. *Perception & Psychophysics, 50*, 93–107. doi.org/10.3758/BF03212211

Kuhl, P.K. (1992). Psychoacoustics and speech perception: internal standards, perceptual anchors, and prototypes. In L.A. Werner, & E.W. Rubel (Eds), *Developmental psychoacoustics* (pp. 293-332). Washington, DC: American Psychological Association.

Kuhl, P.K. (1994). Learning and representation in speech and language. *Current Opinion in Neurobiology, 4*(6), 812-822. https://doi.org/10.1016/0959-4388(94)90128-7

Kuhl, P.K., Conboy, B.T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 979–1000. https://doi.org/10.1098/rstb.2007.2154

Kuhl, P.K., Stevens, E., Hayashi, A., Deguchi, T., Kiritani, S., & Iverson, P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. *Developmental Science*, *9*(2), F13–F21. https://doi.org/10.1111/j.1467-7687.2006.00468.x

Kuhl, P., Williams, K., Lacerda, F., Stevens, K., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*(5044), 606–608. https://doi.org/10.1126/science.1736364

Kutas, M., & Donchin, E. (1977). The effect of handedness, the responding hand and response force on the contra lateral dominance of the "readiness" potential. In J. Desmedt (Ed.), *Attention, voluntary contraction and event-related cerebral potentials* (pp. 189–210). Basel: S. Karger.

Kutas, M., Federmeier K.D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science, 4*(12), 463-470. doi: 10.1016/s1364-6613(00)01560-6. PMID: 11115760.

Lahiri, A., & Reetz, H. (2002). Underspecified recognition. *Laboratory phonology*, *7*, 637–676.

Lahiri, A., & Reetz, H. (2010). Distinctive features: Phonological underspecification in representation and processing. *Journal of Phonetics*, 38(1), 44–59. doi:10.1016/j.wocn.2010.01.002

Lahiri, A. (2018). Predicting universal phonological features. In L. Hyman & F. Plank (Eds.), *Phonological typology*. Berlin: Mouton de Gruyter. pp 229–272.

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, 863. doi:10.3389/fpsyg.2013.00863

Levelt, C. (1994). On the Acquisition of Place. Doctoral dissertation, Leiden University, Leiden, The Netherlands.

Lieberman, P. (1971). Uniquely human. The evolution of speech, thought and selfless behavior. Cambridge (MA), London : Harvard University Press.

Lindblom, B. & Engstrand, O. (1989). In what sense is speech quantal? *Journal of Phonetics, 17,* 107-121.

Ma, W., Golinkoff, R., Houston, D., & Hirsh-Pasek, K. (2011). Word Learning in Infant- and Adult-Directed Speech. *Language Learning and Development, 7,* 209-225. doi: 10.1080/15475441.2011.579839.

Maddieson, I. (1984). *Patterns of Sound,* Campridge: Cambridge University Press.

Mani, N. and Plunkett, K. (2008), Fourteen-month-olds pay attention to vowels in novel words. *Developmental Science*, *11*, 53-59. https://doi.org/10.1111/j.1467-7687.2007.00645.x

Mani, N. & Plunkett, K. (2010). In the Infant's Mind's Ear: Evidence for Implicit Naming in 18-Month-Olds. *Psychological science, 21*, 908-13. doi:10.1177/0956797610373371.

Mani N, & Plunkett K. (2011) Phonological priming and cohort effects in toddlers. *Cognition, 121*(2), 196-206. doi: 10.1016/j.cognition.2011.06.013.

Mann, V.A., & Repp B.H. (1980). Influence of vocalic context on perception of the [ʃ]-[s] distinction. *Perception & Psychophysics, 28*, 213–228. doi: 10.3758/bf03204377.

Markman, E.M. (1990). Constraints children place on word meanings. *Cognitive Science, 14*, 57-77.

Masapollo, M., Polka, L., & Ménard L. (2017). A universal bias in adult vowel perception - By ear or by eye. *Cognition*, *166*, 358–370. doi: 10.1016/j.cognition.2017.06.001

Masapollo, M., Polka, L., Molnar, M. (2017). Directional asymmetries reveal a universal bias in adult vowel perception. *The Journal of the Acoustical Society of America1, 41*, 2857. doi: 10.1121/1.4981006

Maye, J., Werker, J.F. & Gerken, L. (2002). Infant Sensitivity to Distributional Information Can Affect Phonetic Discrimination. *Cognition*, 82, B101-11. 10.1016/S0010-0277(01)00157-3.

Mayor, J., & Plunkett, K. (2009). Using TRACE to Model Infant Sensitivity to Vowel and Consonant Mispronunciations. In N.Taatgen, & H. van Rijn. *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp 1816-1821). Austin, TX: Cognitive Science Society.

Mayor, J., & Plunkett, K. (2014). Infant word recognition: Insights from TRACE simulations. *Journal of Memory and Language, 71*, 89-123.

McClelland, J.L., & Elman, J.L. (1986). The TRACE model of speech perception. *Cognitive Psychology, 18,* 1-86.

Meints, K. & Woodford, A. (2008). Lincoln Infant Lab Package 1.0: A new programme package for IPL, Preferential Listening, Habituation and Eyetracking. http://www.lincoln.ac.uk/psychology/babylab.htm.

Moulton, W. (1962). The Vowels of Dutch: Phonetic and Distributional Classes, *Lingua, 11*, 294-312.

Näätänen, R., & Alho, K. (1997). Mismatch negativity (MMN): The measure for central sound representation accuracy, *Audiology and Neuro-Otology, 2*, 341–353.

Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). *Psychophysiology*, *38*, 1–21.

Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing: A review. *Clinical Neurophysiology, 118*(12), 2544–2590. https://doi.org/10.1016/j.clinph.2007.04.026

Obleser, J., Elbert, T., Lahiri, A., & Eulitz, C. (2003). Cortical representation of vowels reflects acoustic dissimilarity determined by format frequencies. *Cognitive Brain Research, 15*, 207–213.

Pater, J., Stager, C., & Werker, J. (2004). The Perceptual Acquisition of Phonological Contrasts. *Language, 80*, 384-402. 10.1353/lan.2004.0141.

Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., McGinnis, M., & Roberts, T. (2000). Auditory Cortex Accesses Phonological Categories: An MEG Mismatch Study. *Journal of Cognitive Neuroscience*, *12*(6), 1038–1055. https://doi.org/10.1162/08989290051137567

Pierrehumbert, J.B. (2002). Word-specific phonetics. *Laboratory Phonology, 7,* 101–139.

Pierrehumbert, J.B. (2016). Phonological representation: Beyond abstract versus episodic. *Annual Review of Linguistic, 2*, 33-52.

Pliatsikas, C., Wheeldon, L., Lahiri, A., & Hansen, P.C. (2014). Processing of zero-derived words in English: an fMRI investigation. *Neuropsychologia, 53*, 47-53. doi: 10.1016/j.neuropsychologia.2013.11.003. PMID: 24246693.

Poeppel, D., Yellin, E., Phillips, C., Roberts, T.P., Rowley, H. A., Wexler, K., & Marantz, A. (1996). Task-induced asymmetry of the auditory evoked M100 neuromagnetic field elicited by speech sounds. *Cognitive Brain Research, 4*, 231–242.

Polka, L., & Bohn, O.-S. (2003). Asymmetries in vowel perception. *Speech Communication, 41*(1), 221–231. https://doi.org/10.1016/S0167-6393(02)00105-X

Polka, L., & Bohn, O.-S. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of the Acoustical Society of America*, *100*, 577–592.

Polka, L., & Bohn, O.-S. (2011). Natural referent vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics, 39*(4), 467–478. https://doi.org/10.1016/j.wocn.2010.08.007

Polka, L., & Werker, J.F. (1994). Developmental changes in the perception of non-native vowel contrasts. *Journal of Experimental Psychology of Human Perception and Performance, 20*, 421–435.

Post, B., Marslen-Wilson, W.D., Randall, B., & Tyler, L.K. (2008). The processing of English regular inflections: Phonological cues to morphological structure. *Cognition* 109(1), 1–17. https://doi.org/10.1016/j.cognition.2008.06.011

Quam, C., & Swingley, D. (2010). Phonological knowledge guides 2-year-olds' and adults' interpretation of salient pitch contours in word learning. *J. Mem. Lang*, *62*, 135–150. doi: 10.1016/j.jml.2009.09.003

Ramachers, S., Brouwer, S., & Fikkert, P. (2017). How Native Prosody Affects Pitch Processing during Word Learning in Limburgian and Dutch Toddlers and Adults. *Frontiers in Psychology, 8*, 1652. https://doi.org/10.3389/fpsyg.2017.01652

Reetz, H. (2014). Reelex – Reetz–CELEX interface (Version 0.4.5). Retrieved from http://www.phonetik.uni-frankfurt.de/simplex.html

Rice, K. (1995). On vowel place features. *Toronto Working Papers in Linguistics*, *14*. Retrieved from https://twpl.library.utoronto.ca/index.php/twpl/article/view/6321

Rice, K., & Avery, P. (1995). Variability in a deterministic model of language acquisition: A theory on segmental acquisition. In J. Archibald (Ed.), *Phonological acquisition and phonological theory* (pp. 23–42). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rietveld, T., Kerkhoff, J., & Gussenhoven, C. (2004). Word prosodic structure and vowel duration in Dutch. *Journal of Phonetics*, *32*. 349-371. 10.1016/j.wocn.2003.08.002.

Roberts, T.P., Ferrari, P., Stufflebeam, S. M., & Poeppel, D. (2000). Latency of the auditory evoked neuromagnetic field components: Stimulus dependence and insights toward perception. *Journal of Clinical Neurophysiology, 17*, 114–129.

Roberts, A.C., Kotzor, S., Wetterlin, A., & Lahiri, A. (2014). Asymmetric processing of durational differences – Electrophysiological investigations in Bengali. *Neuropsychologia*, *58*, 88–98. https://doi.org/10.1016/j.neuropsychologia.2014.03.015

Sagey, E. (1986). The representation of features and relations in nonlinear phonology. PhD dissertation, MIT.

Schafer, G., & Plunkett, K. (1998). Rapid word learning by fifteen-month-olds under tightly controlled conditions. *Child Development, 69*(2), 309–320. https://doi.org/10.2307/1132166

Scharinger, M. (2009). Minimal representations of alternating vowels*. Lingua*, *119*, 1414–1425.

Scharinger, M., & Idsardi, W. J. (2014). Sparseness of vowel category structure: Evidence from English dialect comparison. *Lingua, 140*, 35–51. https://doi.org/10.1016/j.lingua.2013.11.007

Scharinger, M., & Lahiri, A. (2010). Height Differences in English Dialects: Consequences for Processing and Representation. *Language and Speech*, *53*(2), 245–272. https://doi.org/10.1177/0023830909357154

Scharinger, M., Lahiri, A., & Eulitz, C. (2010). Mismatch negativity effects of alternating vowels in morphologically complex word forms. *Journal of Neurolinguistics, 23*(4), 383–399. https://doi.org/10.1016/j.jneuroling.2010.02.005

Scharinger, M., Monahan, P. J., & Idsardi, W. J. (2012). Asymmetries in the Processing of Vowel Height. *Journal of Speech, Language, and Hearing Research, 55*(3), 903–918. https://doi.org/10.1044/1092-4388(2011/11-0065)

Schlichting, J.E.P.T., & Lutje Spelberg, H.C. (2002). Lexilijst Nederlands: een instrument om de taalontwikkeling te onderzoeken bij Nederlandstalige kinderen van 15-27 maanden in het kader van de vroegtijdige onderkenning. Lisse: Swets Test Publishers.

Schwartz, J.-L., Boe, L.-J., Vallee, N., & Abry, C. (1997). The Dispersion-Focalization Theory of vowel systems. *Journal of Phonetics, 25*, 255-286.

Schwartz, J.-L., Abry, C., Boë, L.-J., Ménard, L., & Vallée, N. (2005). Asymmetries in vowel perception, in the context of the Dispersion-Focalisation Theory. *Speech Communication, 45*(4), 425–434. https://doi.org/10.1016/j.specom.2004.12.001

Schwartz, J.-L., & Escudier, P. (1989). A strong evidence for the existence of a large-scale integrated spectral representation in vowel perception. *Speech Communication, 8*(3), 235–259. https://doi.org/10.1016/0167-6393(89)90004-6

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. https://doi:10.1017/S0142716400001417

Singh, L., Hui, T. J., Chan, C., & Golinkoff, R. (2014). Influences of vowel and tone variation on emergent word knowledge: a cross-linguistic investigation. *Developmental Science,* 17, 94–109. doi: 10.1111/desc.12097

Singh, L., & Quam, C. (2016). Can bilingual children turn one language off? Evidence from perceptual switching. *J. Exp. Child Psychol., 147*, 111–125. doi: 10.1016/j.jecp.2016.03.006

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388*(6640), 381–382. https://doi.org/10.1038/41102

Strauss, T.J., Harris, H.D., & Magnuson, J.S. (2007). jTRACE: A reimplementation and extension of the TRACE model of speech perception and spoken word recognition. *Behavior Research methods, 39*(1), 19-30. DOI: 10.3758/BF03192840.

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition, 76*(2), 147–166. https://doi.org/10.1016/S0010-0277(00)00081-0

Swingley, D., & Aslin, R.N. (2002). Lexical Neighborhoods and the Word-Form Representations of 14-Month-Olds. *Psychological Science,* 13(5), 480–484. https://doi.org/10.1111/1467-9280.00485

Swoboda, P.J., Kass, J., Morse, P.A., & Leavitt, L.A. (1978). Memory factors in vowel discrimination of normal and at-risk infants. *Child Development,* 49(2), 332–339. https://doi.org/10.2307/1128695

Tincoff, R., & Jusczyk, P. W. (1999). Some Beginnings of Word Comprehension in 6-Month-Olds. *Psychological Science, 10*(2), 172–175. https://doi.org/10.1111/1467-9280.00127

Tsao, F., Liu, H., & Kuhl, P.K. (2006). Perception of native and non-native affricate-fricative contrasts: cross-language tests on adults and infants. *The Journal of the Acoustical Society of America, 120*, 2285-2294, https://doi.org/10.1121/1.2338290

Tsuji, S., & Cristia, A. (2014). Perceptual attunement in vowels: A meta-analysis: Infant Vowel Attunement: A Meta-Analysis. *Developmental Psychobiology, 56*(2), 179–191. https://doi.org/10.1002/dev.21179

Tsuji, S., Fikkert, P., Yamane, N., & Mazuka, R. (2016). Language-general biases and language-specific experience contribute to phonological detail in toddlers' word representations. *Developmental Psychology*, *52*(3), 379–390. https://doi.org/10.1037/dev0000093

Tsushima, T., Takizawa, O., Sasaki, M., Shiraki, S., Nishi, K., Kohno, M., Menyuk, P., & Best, C. (1994). Discrimination of English /r-l/ and /w-y/ by Japanese infants at 6-12 months: Language-specific developmental changes in speech perception abilities. *The Emergence of Human Cognition and Language*, 3, 1695-1698.

Twaddell, W. (1938). A note on Old High German Umlaut. *Monatshefte Für Deutschen Unterricht* 30(3/4), 177–181.

Tyler, M., Best, C., Faber, A., & Levitt, A. (2014). Perceptual Assimilation and Discrimination of Non-Native Vowel Contrasts. *Phonetica, 71*, 4-21. 10.1159/000356237.

Van der Feest, S.V.H. (2007). Building a phonological lexicon: the acquisition of the Dutch voicing contrast in perception and production. PhD dissertation, Radboud University Nijmegen.

Van der Feest, S.V.H., & P. Fikkert (2015). Building phonological lexical representations. *Phonology,* 32, 207–239. https://doi:10.1017/S0952675715000135

Van der Feest, S.V.H., Fikkert, P., & Davis, B.L. (2016). Cross-linguistic Differences in the Perception of Dorsal and Coronal CV-Combinations: Evidence from English and Dutch. In: *Proceedings of the Annual Boston University Conference on Language Development Volume 2* (pp. 417-428). Cascadilla Press, Somerville, MA.

Van Oostendorp, M. (1995). Vowel Quality and Phonological Projection. PhD Thesis, Tilburg University.

Werker, J.F., Fennell, C.T., Corcoran, K.M., & Stager, C.L. (2002). Infants' Ability to Learn Phonetically Similar Words: Effects of Age and Vocabulary Size. *Infancy, 3*(1), 1-30.

Werker, J.F., & Fennell, C.T. (2004). Listening to Sounds versus Listening to Words: Early Steps in Word Learning. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 79–109). MIT Press.

Werker, J.F., & Tees, R.C. (1984). Cross-Language Speech Perception: Evidence for Perceptual Reorganization During the First Year of Life. *Infant Behavior and Development, 7*, 15.

Werker, J.F., & Tees, R.C. (1999). Influences on infant speech processing: Toward a new synthesis. *Annual Review of Psychology, 50*, 509–535. https://doi.org/10.1146/annurev.psych.50.1.509

White, K.S., & Morgan, J.L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*, 114-132.

Wiese, R. (1996). Phonological versus morphological rules: On German Umlaut and Ablaut. *Journal of Linguistics, 32*(1), 113–135. doi:10.1017/S0022226700000785

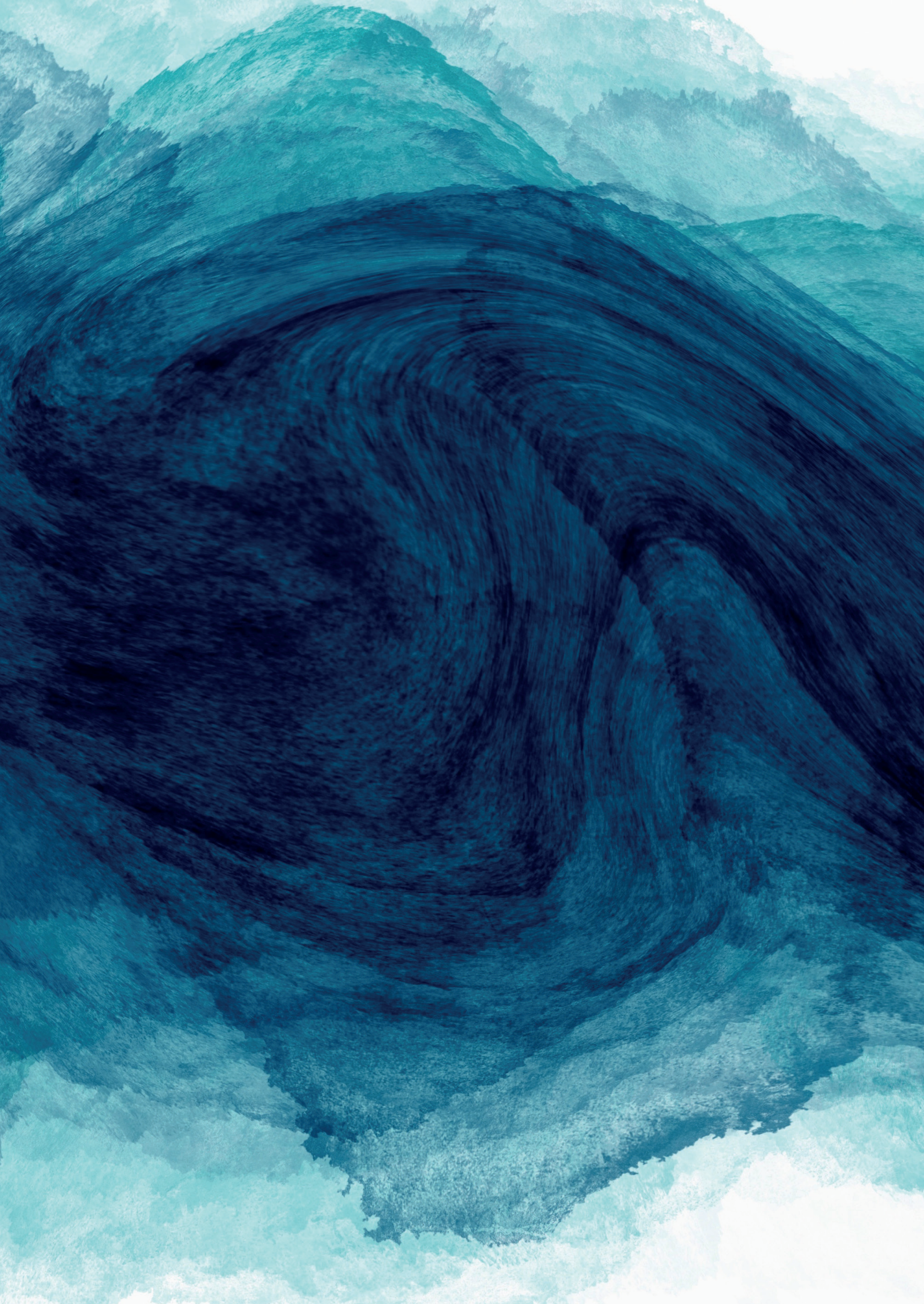
Winkler, I., Lehtokoski, A., Alku, P., Vainio, M., Czigler, I., Csepe, V., Aaltonen, O., Raimo, I., Alho, K., Lang, H., Iivonen, A., & Näätänen, R. (1999). Pre-attentive detection of vowel contrasts utilizes both phonetic and auditory memory representations. *Cognitive Brain Research*, *7*, 357–369.

Zink, I., & Lejaegere, M. (2002). *N-CDI's: Lijsten voor Communicatieve Ontwikkeling. Aanpassing en hernormering van de MacArthur CDI's van Fenson et al.* Leuven/Leusden: Acco.

Zink, I., & Lejaegere, M. (2007). *N-CDI-3, aanpassing en herhormering van de MacArthur CDI: Level III van Dale et al.* (2000). Leuven/Voorburg: Acco

Zonneveld W., & Trommelen M. (1980) Egg, Onion, Ouch! On The Representation Of Dutch Diphthongs. In: Geerts G. et al. (eds) *Dutch Studies*. Springer, Dordrecht. https://doi.org/10.1007/978-94-009-8855-2_13

Zwaardemaker, H., & Eijkman, L.P.H. (1928). Leerboek der Phonetiek: inzonderheid met betrekking tot het Standaard-Nederlandsch. Haarlem: Erven F. Bohn.



# Appendices

## APPENDIX 1: Stimuli creation

Here we provide a brief description of how the stimuli used in Chapter 2 were created. Before using the stimuli in the EEG experiment, three pretests (see Appendix 2) were performed to evaluate and adjust the stimuli.

### Recordings

A male speaker (age 28) produced the three different vowels (/e/, /o/, and /ø/) in a set (i.e. triad) in both /V/ (where V stands for *vowel*) and /hV/ context (e.g. /e, o, ø/, or /he, ho, hø/). A difference in context does not make a difference for the vowels' place of articulation features. The voiced consonant /h/ is articulated with glottal friction, with the same shape of vocal tract as that of the following vowel (e.g. Booij, 1995). Already during pronunciation of the /h/, the oral cavity adopts the shape to produce the subsequent V.

The speaker was instructed to pronounce all vowels with similar and steady pitch, and similar duration. We used triads to facilitate similar pitch, duration and loudness features for the different vowels. However, when our speaker pronounced a triad, he would produce the first syllable with rising pitch, and the third syllable with falling pitch. Even after extensive and explicit instruction this list intonation still appeared in his speech. We resolved this issue by having him produce triads in which the order of vowels differed, (e.g. /e o ø/ /o ø e/, and /ø e o/ and similarly for hV context). The syllable in the middle received a relatively constant pitch, and all three vowels appeared in this position.

### Stimuli preprocessing

All preprocessing steps were conducted in PRAAT (version 5.3.22). Tokens of the recorded vowels were selected based on similarity in pitch (within and between category) and formant contours (within category), resulting in three most similar tokens of each vowel category. Eventually, all tokens that were selected for the experiment of /o/, and /ø/ originated from a hV syllable, while all /e/ tokens originated from a V context.

Selected vowels were cut to have a 200 ms duration. As vowels' features at their centre are most representative for their specific vowel category (Rietveld & van Heuven, 2001), the midpoint relative to the duration of the original vowel served as the midpoint for extraction of the 200 ms portion. For all vowels, 200 ms was extracted; for none of the stimuli original vowel initiation was preserved. Before determining the midpoint of the vowel, /h/ was removed for the /o/ and /ø/ tokens. For hV syllables, /h/ was removed before further steps were carried out. Of the resulting tokens, mean amplitude was scaled at 70 dB, and amplitude ramps of 50 ms were applied on both onset and offset using the Audio Manipulation script by Kerkhoff (2009).

## APPENDIX 2: Pretests of stimuli

Before using the stimuli the EEG experiment in Chapter 2, their quality was determined in three behavioural pretests.

- *Pretest 1* assessed whether differences in original vowel context (either /hV/ or /V/ context) was still audible. To use the stimuli in subsequent EEG experiments, no audible difference between contexts was required.
- In *Pretest 2* we controlled for perceptual loudness differences for the three different vowel categories. Perceptual loudness differences might occur as a result of differences in characteristic formant frequencies for the different categories.
- In *Pretest 3* we evaluated perceptual variability between the tokens within each vowel category. To use the stimuli in subsequent EEG experiments, no difference in within-category variation should be perceived.

**Participants**

Pretest 1 and Pretest 2 were performed by two adult male subjects (both age 28) and were performed in a quiet room. Both subjects participated on voluntary basis. Pretest 3 was performed by 16 adult participants (50% females) (ages ranging between 20 and 29; M = 24.6; SD = 3.3) in a sound attenuated booth. Participants of Pretest 3 were reimbursed for their participation: 5 Euro or 0,5 study credits. All participants in all pretests were right-handed adult native Dutch listeners, consistent with the participants eligible for participating in the eventual EEG experiment. They had no background in linguistics, and reported normal hearing, and no speech or language impediments. No dialect speakers were included.

**PRETEST 1: Original context: /hV/ vs. /V/**

In Pretest 1 we assessed whether the original context of the stimuli was still audible. All tokens of /o/ and /ø/ originated from an hV context, while all /e/ tokens had a V context. This difference in original context should not be detectible in our stimuli. To verify this, we randomly presented all 9 individual stimuli on a laptop (HP Probook 6750B) using Presentation Software (version 18.2 02.18.16, www.neurobs.com). All stimuli were presented 3 times, resulting in 27 trials. Subjects judged whether they heard /V/ or /hV/, responding by respectively pressing keys [1] or [2] on the keyboard. Stimuli were auditorily presented through over ear headphones (Sennheiser HD 215) at fixed volume.

Participants were unable to detect context above chance level. Participants reported that all trials just seemed to be /V/, and they did not hear the initial /h/. In 33% of /o/ and /ø/ trials participants selected /hV/ context, which is considerably below chance level of 50%. As no obvious /hV/ stimuli were presented, participants became motivated to press /hV/ also in trials where they heard no obvious /h/. This has probably fed /hV/ responses. Participants were not consistent in their judgement, selecting /hV/ at one trial, and /V/ at another trial for identical stimuli. We concluded that, as we expected, the initial /h/ was successfully deleted from the signal as /h/ was no longer perceivable.

**PRETEST 2: Perceptual loudness**

Due to non-linear frequency response characteristics of the human auditory system, different frequencies with matched amplitude (objective loudness in dB) do not necessarily have similar perceptual loudness. For example, low frequency sounds are perceptually louder than high frequency sounds. These effects differ at different intensities. The stimuli are to be used in Mismatch Negativity EEG experiments. To avoid mismatch negativity responses driven by differences in perceptual loudness between the three vowel categories of /e/, /o/ and /ø/ in the EEG experiments, we controlled for perceptual loudness following the procedure described below. In Pretest 2 we assessed whether the stimuli were similar in perceptual loudness. Different vowel categories have different frequency characteristics.

In PRAAT (version 5.3.35; Boersma & Weenink, 2012) we normalized all stimuli to have a matching RMS amplitude of 70 dB. We ran a judgement task on a laptop (HP Probook 6750B) using Presentation Software. In this AX-task, all possible stimulus pairs were presented to the participants. Stimuli were presented through headphones (Sennheiser HD 215). Each stimulus was compared to each of the other stimuli (resulting in within-category trials and between-category trials), but never to itself, thus resulting in 9*8 = 72 trials, presented in random order. Subjects judged whether stimulus X (= the second stimulus of a trial) was less loud, equally loud, or louder than stimulus A by choosing respectively keys [1], [2] or [3].

Subjects reported that /e/ stimuli sounded louder than the other two vowel categories. The button-press data support this judgement. Surprisingly, this perceptual difference appeared asymmetric. In 71,9% of trials with /e/ presented last (stimulus X) and another vowel category as stimulus first (stimulus A), /e/ was considered louder. However, in the reversed situation where A = /e/ and X was any token of the other two vowel categories, /e/ was considered louder in only 9,4% of the trials. This implies a recency effect. A difference in perceptual loudness can therefore be considered small. In only 2,7% (= 4 cases), a non-/e/ token was considered louder than another token. Only once (<1%) an /e/ sound was considered softer than a stimulus of the other two vowel categories.

Relative loudness of sounds in air as perceived by the human ear are expressed in dBA. Perceptual loudness of all stimuli was measured using a dB(A) meter, measuring loudness as produced by headphones. Consistent with the results of the judgement task, the /e/ sounds resulted in slightly higher dBA values (+/- 67 dBA for /e/ and +/- 64 dBA for the other two categories) despite identical dB. Hence, amplitude of /e/-stimuli was attenuated in PRAAT until dBA-values of /e/ tokens were similar to the other two vowel categories. As a result, all tokens of all categories were within a 2 dBA range with a mean of 64 dBA. In all following experiments, volume was set to achieve this same dBA intensity, thus controlling for perceptual loudness.

**PRETEST 3: Within-category variation**

After the intensity adjustments as mentioned in Pretest 2 we proceeded with a more thorough pretest. In this third pretest we assessed whether all three vowel categories comprise similar degrees of within-category variation.

By using several natural tokens of each vowel, acoustic within-category variability is introduced. This simulates more natural speech perception conditions and – most importantly – forces a listener's processing system to map the incoming acoustic signals onto more abstract representations (Eulitz & Lahiri, 2004). In other words, subjects will listen in a categorical way rather than focus on minor acoustic differences between tokens. Thus, some within-category variation is desired. However, for each category (/o/, /e/ and /ø/) we aim for similar degrees of within-category variability. If there is a particular token within a category that stands out, this could lead to within-category elicitation of a Mismatch Negativity response and could thus have an undesirable influence on the overall category MMN response. Differences in MMN for the different vowel categories should not be due to differences in within-category variation. Thus, in Pretest 3, we tested whether the perceptual variation within the vowel categories was similar for all three vowel categories.

We performed an ABX task (or *matching-to-sample* task); a discrimination procedure where triads of stimuli are presented. The first two (A and B) are references and the third/last stimulus (X) is compared to these references. Subjects were instructed to determine whether the third vowel (stimulus X) in a triad was more similar to either the first (stimulus A) or the second (stimulus B) sound. They responded by pressing button [1] for stimulus A or [2] for stimulus B on the computer's keyboard. Stimuli were presented through headphones in a sound attenuated booth. Volume was set at 64 dBA for all stimuli. Participants were not allowed to adjust the volume. Inter-stimulus-interval (ISI) was set at 700 ms, similar to the eventual EEG experiments. Before the start of the experiment participants were informed about the experiment and signed a consent form.

The ABX-experiment was run using Presentation Software. Within each block, trials were presented in random order. The experiment started with 6 practice trials with between-category trials. Then, 3 test blocks of 54 within-category trials (e.g. A = [e1] - B = [e3] - X = [e2]) were presented (PART 1). Each individual triad occurred in each of the three blocks. In addition, a final test block (PART 2) was included (36 trials), containing only between-category trials (e.g. A = [e1] - B = [o2] - X = [o3]). The three blocks of within-category trials (PART 1) can show the degree of within-category variation. PART 2, testing between-category trials, can provide proof that participants can behaviourally discriminate the vowels of different categories. In between blocks subjects were free to take breaks. After the experiment, subjects filled out a post-test questionnaire regarding strategy use, difficulty, and assumptions with respect to research questions.

Only responses with *100 > RT < (average RT + 2SD)* were included in the data assessment.

Results showed that participants were able to discriminate between vowel categories /e/, /o/, and /ø/ (92,5% correct) (PART 2 in the experiment).

In PART 1 we can distinguish two conditions: (1) *Correctness_Condition*: condition where X is identical to A or B and a correct/incorrect answer can be defined, and (2) *Preference_Condition*: condition where X is not identical to either A or B. Thus, no correctness can be assessed. The latter condition provides information about the degree of similarity of X to A or B and thus provides info about the within-category variation. As we expected, identifying a stimulus in a within-category trial (Correctness Condition in PART 1) was close to chance (56,6% correct). In the Preference Condition, option A was selected 51% of the time as opposed to 49% for option B. This is a very close approximation of random ratio. Within each vowel category, random responses would lead to 1/3 of responses for each token. Results were close to this ratio (see Table A2.1). These ratio's did not deviate enough from 33,3% to conclude that there were salient differences between the tokens within each category.

**Table A2.1** Response ratio's to each token of each vowel category in PART 1 of the ABX-pretest.

| | /e/ | /o/ | /ø/ |
|---|---|---|---|
| token 1 | 36% | 30% | 36% |
| token 2 | 23% | 36% | 30% |
| token 3 | 41% | 34% | 34% |

## APPENDIX 3: Stimuli characteristics

**Table A3.1** Fundamental frequencies (Hz) of all tokens used as stimuli.

| Vowel | F0 Token 1 | F0 Token 2 | F0 Token 3 | Mean F0 |
|---|---|---|---|---|
| /e/ | 111.8 | 112.7 | 113.0 | 112.5 |
| /ø/ | 113.4 | 114.3 | 114.1 | 113.9 |
| /o/ | 112.0 | 113.2 | 112.8 | 112.7 |

**Table A3.2** Formant frequencies (Hz) for formants F1, F2 and F3 of three tokens of /e/. Measures sampled at 25%, 50% and 75% of vowel.

| | **Token [e]1** | | | **Token [e]2** | | | **Token [e]3** | | | **Adank et al. /e/** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***t*** | **25%** | **50%** | **75%** | **25%** | **50%** | **75%** | **25%** | **50%** | **75%** | **25%** | **50%** | **75%** |
| **F1** | 517 | 458 | 336 | 415 | 400 | 361 | 472 | 472 | 354 | 541 | 348 | 303 |
| **F2** | 1822 | 1881 | 1951 | 1887 | 1995 | 2075 | 1847 | 1877 | 1928 | 1901 | 1977 | 2013 |
| **F3** | 2230 | 2341 | 2482 | 2545 | 2583 | 2597 | 2254 | 2347 | 2498 | 2416 | 2518 | 2600 |

The averaged values for formant frequencies are taken at the 20-30% interval (t=25%), 45-55% (t=50%) interval and 70-80% (t=75%) interval. t = time point. Each vowel had a duration of 200 ms. As a reference, average formant frequencies as reported in Adank, van Hout, and Smits (2004) for Northern standard Dutch by male speakers are included.

**Table A3.3** Formant frequencies (Hz) for F1, F2 and F3 of three tokens of /ø/.

| | **Token [ø]1** | | | **Token [ø]2** | | | **Token [ø]3** | | | **Adank et al. /ø/** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***t*** | **25%** | **50%** | **75%** | **25%** | **50%** | **75%** | **25%** | **50%** | **75%** | **25%** | **50%** | **75%** |
| **F1** | 384 | 365 | 320 | 407 | 375 | 329 | 409 | 337 | 296 | 396 | 344 | 295 |
| **F2** | 1432 | 1535 | 1607 | 1544 | 1563 | 1592 | 1475 | 1573 | 1588 | 1472 | 1594 | 1591 |
| **F3** | 2123 | 2104 | 2151 | 2254 | 2241 | 2229 | 2126 | 2170 | 2070 | 2125 | 2174 | 2072 |

Measures sampled at 25%, 50% and 75% of vowel. The averaged values for formant frequencies are taken at the 20-30% interval (t=25%), 45-55% (t=50%) interval and 70-80% (t=75%) interval. t = time point. Each vowel had a duration of 200 ms. As a reference, average formant frequencies as reported in Adank, van Hout, and Smits (2004) for Northern standard Dutch by male speakers are included.

**Table A3.4** Formant frequencies (Hz) for F1, F2 and F3 of three tokens of /o/.

| | Token [o]1 | | | Token [o]2 | | | Token [o]3 | | | Adank et al. /o/ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| ***t*** | **25%** | **50%** | **75%** | **25%** | **50%** | **75%** | **25%** | **50%** | **75%** | **25%** | **50%** | **75%** |
| **F1** | 530 | 420 | 294 | 444 | 412 | 365 | 527 | 420 | 289 | 524 | 401 | 252 |
| **F2** | 1013 | 976 | 877 | 1015 | 929 | 848 | 1021 | 990 | 881 | 1017 | 968 | 892 |
| **F3** | 2250 | 2230 | 2286 | 2258 | 2306 | 2371 | 2307 | 2237 | 2261 | 2249 | 2244 | 2272 |

Measures sampled at 25%, 50% and 75% of vowel. t = time point. The averaged values for formant frequencies are taken at the 20-30% interval (t=25%), 45-55% (t=50%) interval and 70-80% (t=75%) interval. Each vowel had a duration of 200 ms. As a reference, average formant frequencies as reported in Adank, van Hout, and Smits (2004) for Northern standard Dutch by male speakers are included.

**Figures A3.1-A3.6** show the formats in the three Dutch and German vowels:

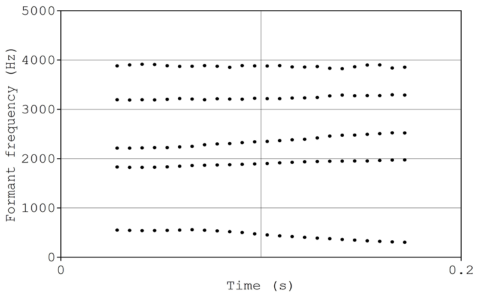



**Figure A3.1.** Vowel formants Dutch /e/

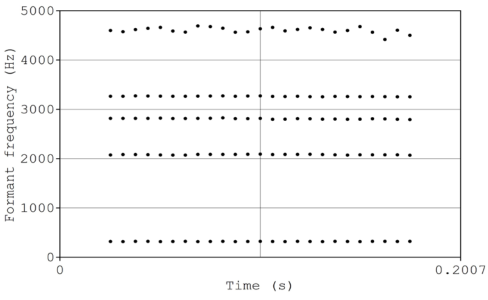



**Figure A3.2.** Vowel formants German /e/

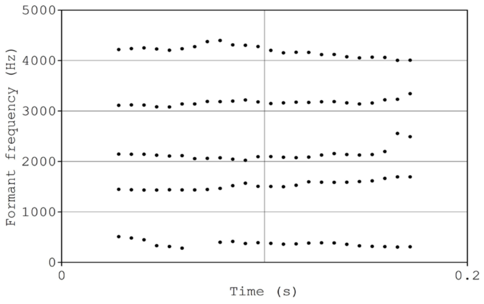



**Figure A3.3.** Vowel formants Dutch /ø/

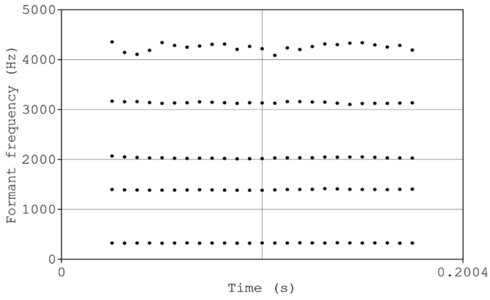



**Figure A3.4.** Vowel formants German /ø/

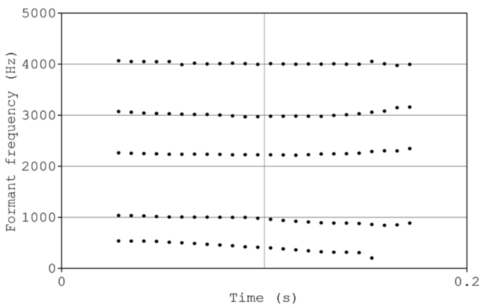



**Figure A3.5.** Vowel formants Dutch /o/

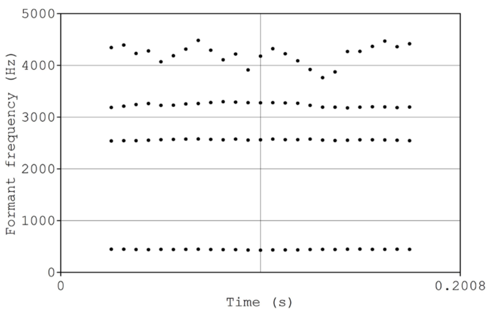



**Figure A3.6.** Vowel formats German /o/

## APPENDIX 4: RMS data

For sake of comparability, similar to Eulitz and Lahiri (2004), we also ran an RMS (root mean square) analysis on the data in Chapter 2.

**Table A4.1.** Results root-mean-square (RMS) MMN amplitude measure for six conditions.

| Vowel pair | Contrastive feature | Condition | RMS amplitude ± SEM (μV) |
|---|---|---|---|
| /ø - o/ | place | [ø]/o/ | 1.15 ± 0.173 |
| | | [o]/ø/ | 0.77 ± 0.096 |
| /e - ø/ | labiality | [e]/ø/ | 0.93 ± 0.149 |
| | | [ø]/e/ | 1.48 ± 0.211 |
| /e - o/ | place & labiality | [e]/o/ | 0.95 ± 0.105 |
| | | [o]/e/ | 1.17 ± 0.137 |

An interaction of *Pair of Inversion * Direction of Change* was significant (F(1/16) = 7.46; p = 0.015. No main effect for *Pair of Inversion* was found: F(1/16) = 3.4; p = 0.08. Also no significant main effect was found for *Direction of Change:* F(1/16) = 0.53; p = 0.47. The found interaction implies that in processing of these vowel pairs, the impact of direction of change differs for different contrasts.

Apart from this ANOVA, we ran planned comparisons. For the place contrast /ø-o/, paired samples t-tests for the difference between [ø]/o/ (M = 1.15; SD = 0.71) and [o]/ø/ (M = 0.77; SD = 0.93) did not give significant results; t(16) = -1.96, p = 0.067.

For the labiality contrast /e-ø/, paired samples t-tests resulted in a significant asymmetry: t(16) = 2.59, p = 0.020, with a larger amplitude for [ø]/e/ (M = 1.48; SD = 0.62) than for [e]/ø/ (M = 0.93; SD = 0.87).

For the two-feature contrast /e-o/, paired samples t-tests for the for the difference between [e]/o/ (M = 0.95; SD = 0.43) and [o]/e/ (M = 1.17; SD = 0.57) did not give significant results; t(16)=-2.004, p = 0.062.

## APPENDIX 5: Familiarity and semantic similarity judgement

To assess the adequacy of our stimuli in Chapter 3, we conducted two pretests. We ran a familiarity judgement task (63 participants) as well as a semantic similarity judgement task (75 participants) as online questionnaires in Qualtrics Software (https://www.qualtrics.com ), Version February 2017. Both pretests are discussed in more detail below.

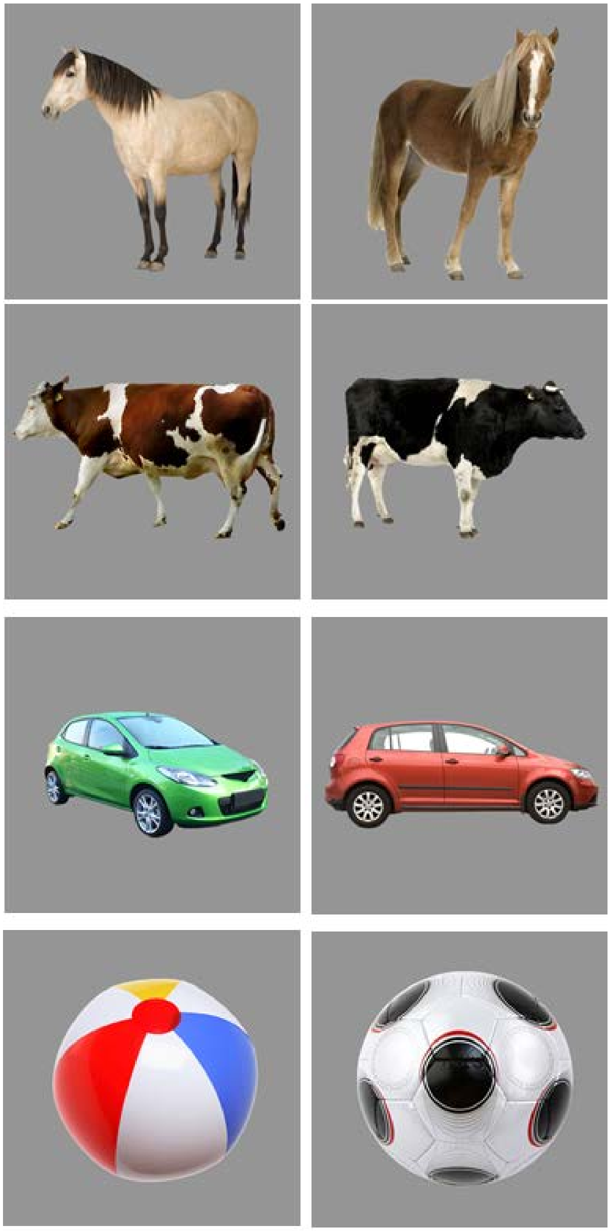
Frequency counts from corpora / linguistic databases such as CELEX or SUBTLEX are a conventional way of getting a sense of how familiar people are with particular words, implying how well subjects know words. Words participants do not know lack semantic content and can never yield semantic priming. However, actual experiences are never identical to counts in a database. Hence, a familiarity judgement task was conducted to assess whether stimuli were familiar enough to be included in the experiment. Familiarity ratings were scored on a 5-point-scale (1 = unfamiliar; 5 = very familiar) similar to Roberts, Kotzor, Wetterlin, and Lahiri (2014). Items were distributed over four different versions of Qualtrics questionnaires. Participants were randomly assigned to a particular version. Participants were instructed to score 1 for words they did not know at all or 2 for words they had heard but were unsure about its meaning. 3 would mean they did know the words meaning well although they did not hear or use it regularly, 4 for the words that they know well and also use quite regularly, and 5 for words they knew well, came across often and used frequently. Only words that scored an mean of 3.5 or more were included in the experiment.

In addition, prime-target pairs were judged in an online semantic similarity judgement task as well, to assess whether they were strong enough synonyms to permit semantic priming – or in the case of word pairs to be used in unrelated conditions, whether they were sufficiently unrelated to prevent semantic priming. Semantic similarity was rated on a 7-point-scale (1 = entirely unrelated in meaning; 7 = full synonyms). Only word pairs that had an average score > 5 were included in the experiment to serve in the semantically related conditions (CP_related & MP_related). An average score < 2 was required to include the items in the unrelated/control conditions (CP_control & MP_control).

Based on both pretests, words and word pairs were changed to improve them prior to conducting the experiment. Adapted or newly created items were included in posttest judgement tasks once again – both a familiarity and a semantic similarity judgement task, in order to eliminate poor items before running final analysis.

## APPENDIX 6: Familiar objects in filler trials

Four familiar words occurred in each test block: horse, cow, car and bal. Pictures used for the filler trials are listed here. Of each object there are two pictures, of which one occurred as a target and one as a distractor in a test block.



## APPENDIX 7: MP-effect based on naming effects

Naming effects calculated based on difference score (i.e. (PTL post-naming) – (PTL pre-naming)). CP = correct pronunciation. MP = mispronunciation. SR = surface representation. SE = standard error. Naming effects were compared to chance (i.e. zero) using t-tests.

One may argue that since we did not find a significant difference in Chapter 4 between CP and MP for two conditions (one only marginally non-significant), perhaps word recognition was not hindered for these conditions and in fact there are perceptual asymmetries. However, comparing CP and MP naming effects means that two data sets are compared which both have a lot of variance and great overlap between distributions. The ample amount of variance means that perhaps the current study may lack power to reveal significant differences between CPs and MPs even when MPs had no naming effect and CPs did have a significant positive naming effect.

Naming effects were calculated as difference scores for each trial for each child and tested in a t-test against chance level. Since in this test only one data set with variance is included, naming effects are easier to detect than differences between CP and MP, where both distributions are overlapping. Thus, naming effects are more robust. Furthermore, comparing CP and MP is most informative in case an MP yields a positive naming effect, since then a CP-MP difference can reveal hindered word recognition nonetheless. Our study, however, did not yield positive naming effects for any of the MPs. This already means that word recognition is unsuccessful. A CP-MP comparison hence does not to the interpretation.

| Vowel pair | Contrastive feature(s) | Distance | CP | MP | CP SR features | MP SR features | CP naming effect | MP naming effect | CP-MP contrast (SE) |
|---|---|---|---|---|---|---|---|---|---|
| /e/~/ø/ | labiality | 1 | /e/ | [ø] | [CORONAL] | [CORONAL] [LABIAL] | **0.21** ** (SD 0.34) (SE 0.02) **<0.0001** | **-0,04** (SD 0.34) (SE 0.04) **p=0.17 (neg)** | 0,255 (0,055) ** |
| | | | /ø/ | [e] | [CORONAL] [LABIAL] | [CORONAL] | **0.13** ** (SD 0.35) (SE 0.02) **<0.0001** | **0.04** (SD 0.37) (SE 0.04) **p=0.13 (pos)** | 0,072 (0,053) p = 0.176 |
| /ø/~/o/ | place | 1 | /ø/ | [o] | [CORONAL] [LABIAL] | [DORSAL] [LABIAL] | **0.13** ** (SD 0.35) (SE 0.02) **<0.0001** | **-0.01** (SD 0.41) (SE 0.05) **p=0.42 (neg)** | 0,131 (0,057) * |
| | | | /o/ | [ø] | [DORSAL] [LABIAL] | [CORONAL] [LABIAL] | **0.14** ** (SD 0.34) (SE 0.02) **<0.0001** | **0.04** (SD 0.43) (SE 0.05) **p=0.20 (pos)** | 0,100 (0,057) p = 0.078 |
| /o/~/e/ | labiality & place | 1 | /o/ | [e] | [DORSAL] [LABIAL] | [CORONAL] | **0.14** ** (SD 0.33) (SE 0.02) **<0.0001** | **-0.05** (SD 0.42) (SE 0.05) **p=0.15 (neg)** | 0,200 (0,054) ** |
| | | | /e/ | [o] | [CORONAL] | [DORSAL] [LABIAL] | **0.21** ** (SD 0.34) (SE 0.02) (SE 0.02) **<0.0001** | **-0.11** * (SD 0.38) (SE 0.04) **p=0.0072 (neg)** | 0,325 (0,055) ** |

## APPENDIX 8: MP-effect based on post-naming PTL

Mean PTL in pre- and post-naming windows in CP and MP trials. MP-effect calculated based on PTL in post-naming window in CP versus MP for that direction of change. Estimated contrast (MP-effect) reported in Least Square Means. ** = p < 0.0001.

| | | CP | | MP | | MP-effect |
|---|---|---|---|---|---|---|
| **Vowel Pair** | **Condition** | **PTL pre-naming (SE)** | **PTL post-naming (SE)** | **PTL pre-naming (SE)** | **PTL post-naming (SE)** | **LSM (SE)** |
| /e/~ /ø/ | /e/[ø] | 0.498 (0.03) | 0.710 (0.03) | 0.528 (0.03) | 0.483 (0.03) | 0.234 (0.04)** |
| | /ø/[e] | 0.509 (0.03) | 0.705 (0.03) | 0.485 (0.03) | 0.535 (0.03) | 0.153 (0.03)** |
| /ø/~ /o/ | /ø/[o] | 0.623 (0.03) | 0.671 (0.03) | 0.499 (0.03) | 0.488 (0.04) | 0.199 (0.04)** |
| | /o/[ø] | 0.528 (0.03) | 0.695 (0.03) | 0.501 (0.03) | 0.548 (0.03) | 0.137 (0.04)** |
| /o/~ /e/ | /e/[o] | 0.514 (0.03) | 0.722 (0.03) | 0.500 (0.03) | 0.383 (0.03) | 0.236 (0.03)** |
| | /o/[e] | 0.551 (0.03) | 0.679 (0.03) | 0.503 (0.03) | 0.450 (0.03) | 0.329 (0.04)** |

For the current analysis we used the same model as reported in the main text of Chapter 4, only using post-naming PTL as our dependent variable instead of the difference score between post- and pre-naming PTL. Current analysis shows a significant MP effect for all directions of change. This supports that children are sensitive to all vowel changes we tested. However, since we found a small but significant target preference in the pre-naming window, which means we cannot base our conclusions on this post-naming PTL rather than on a difference score between pre- and post-naming (i.e. naming effect) is not approved.

## APPENDIX 9: PTL time course graphs

The time course graphs below all reflect PTL throughout the trial in the experiment described in Chapter 4. Calculations include all looks – not only looks that were on the distractor during naming. The vertical dotted line indicates onset of the target word at 2500 ms into the trial. The horizontal solid line at the top of each graphs indicates the time window (3000 – 6000 ms) on which we did our naming effect analysis.

### Place contrast

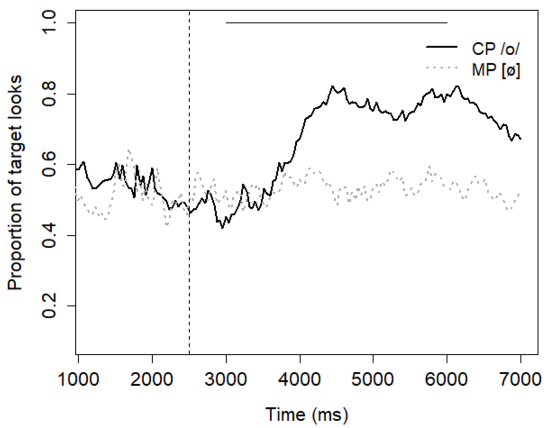



**Figure A8.1** Condition [ø]/o/: change from dorsal CP to coronal MP.

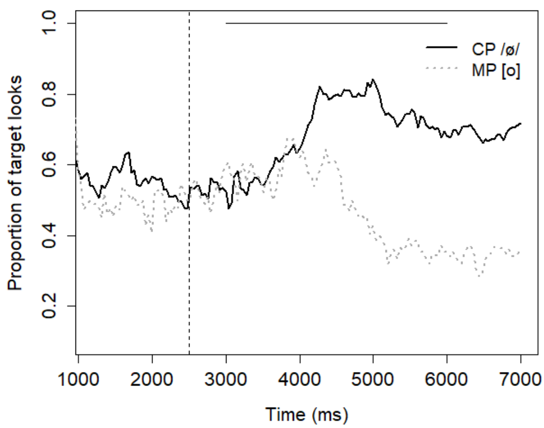



**Figure A8.2** Condition [o]/ø/: change from coronal CP to dorsal MP.

## Lability contrast

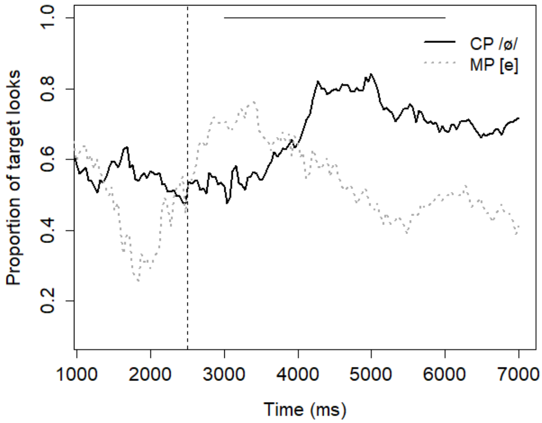



**Figure A8.3** Condition [e]/ø/: change from labial CP to nonlabial MP.

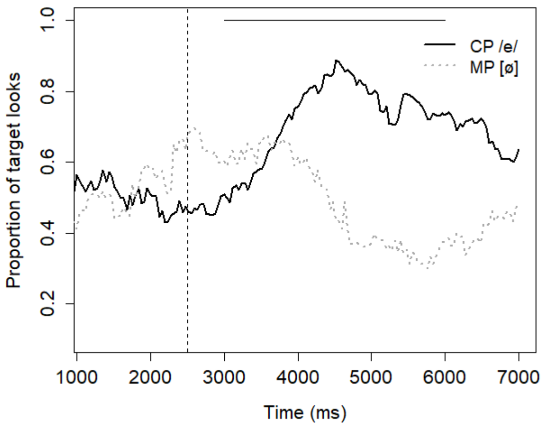



**Figure A8.4** Condition [ø]/e/: change from nonlabial CP to labial MP.

## Two-feature contrast

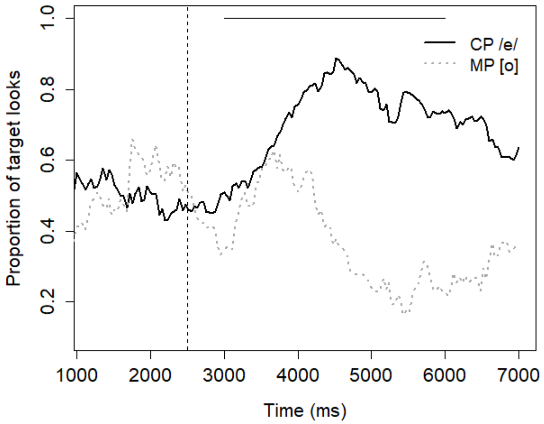



**Figure A8.5** Condition [o]/e/: change from coronal nonlabial CP to dorsal labial MP.

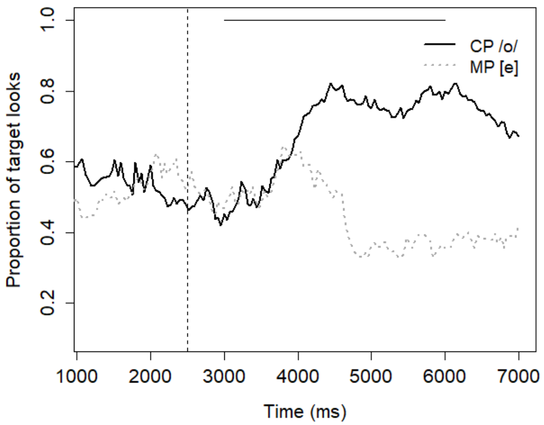
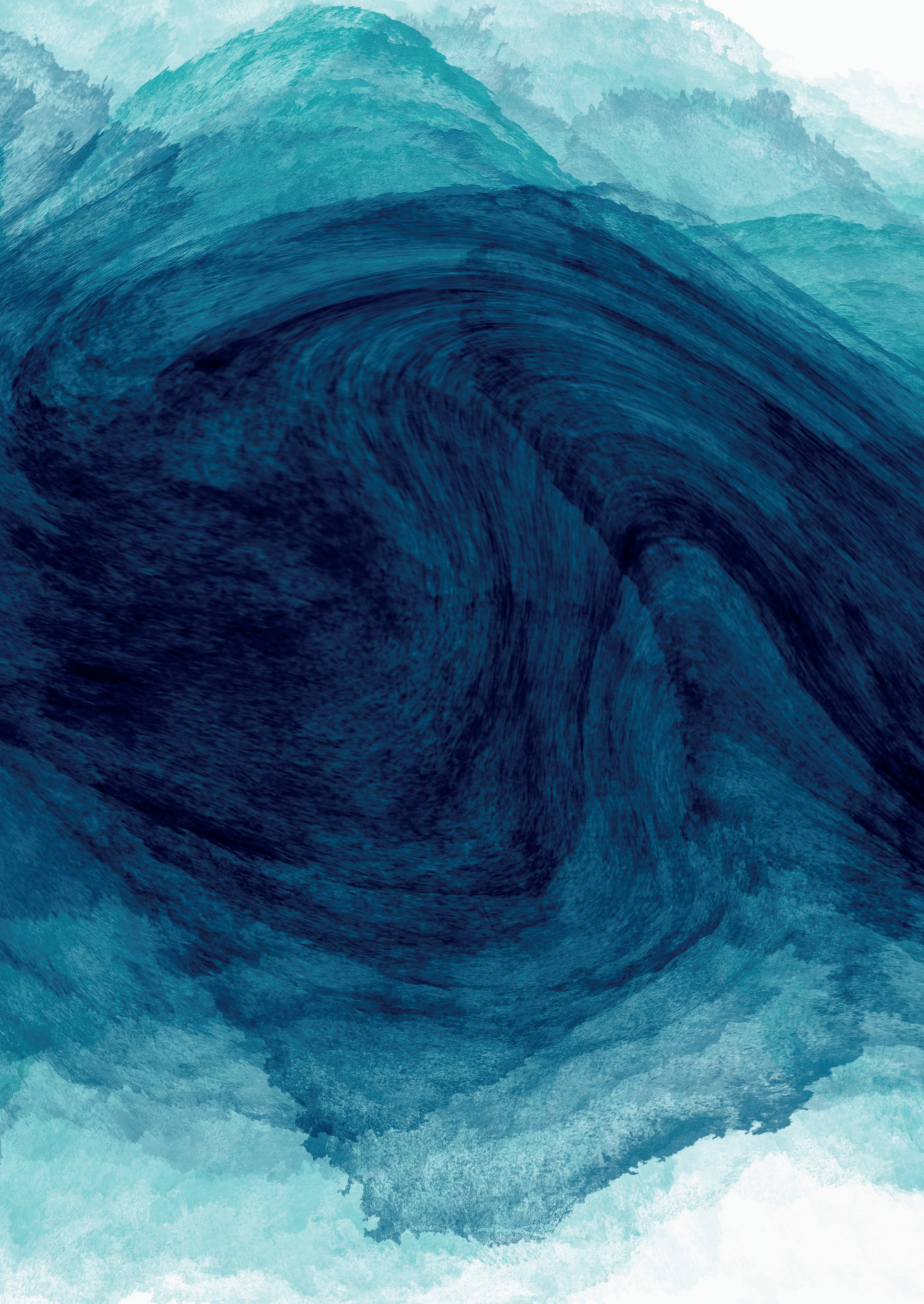



**Figure A8.6** Condition [e]/o/: Change from dorsal labial CP to coronal nonlabial MP.

## APPENDIX 10: Neighbourhood density

Neighbours are defined as nouns that differ by substituting one phoneme, based the Lexilijst Nederlands (Schlichting & Lutje Spelberg, 2002).

| Word | Noº of neighbours | Neighbours |
|---|---|---|
| /tes/ | 1 | tas |
| /tos/ | 2 | doos, tas |
| /tøs/ | 2 | neus tas |
| /den/ | 2 | been, teen |
| /don/ | 1 | dood |
| /døn/ | 1 | deur |



# Dankwoord

Voor je ligt mijn proefschrift. Het is af. Het is er echt. Ik kan het nog niet helemaal geloven. Mijn naam staat erop, maar dit proefschrift is er niet alleen dankzij mijzelf. Er zijn mensen die hebben bijgedragen aan de totstandkoming. Die mensen wil ik bij dezen graag bedanken.

Mijn dank gaat uit naar mijn promotoren Paula Fikkert en Aditi Lahiri en naar mijn copromotor Tineke Snijders voor het samen met mij aangaan van dit avontuur.

Paula, zonder jou had ik nooit een onderzoeksaanvraag geschreven voor het NWO. Na het begeleiden van mijn tweede masterscriptie zei je dat je dacht dat ik kans maakte op een persoonsgebonden beurs. We hebben toen samen een onderzoeksvoorstel uitgedacht. En wonder boven wonder kreeg ik inderdaad de beurs. Dit proefschrift had er anders niet gelegen. Je liet me vrij in het uitvoeren van dit onderzoek. Daar heb ik veel van geleerd. Bedankt voor je begrip en steun wat betreft de gezondheidsproblemen die mij tijdens dit promotietraject parten speelden en die de toch al turbulente weg naar een afgerond proefschrift extra ingewikkeld maakten.

Aditi, de discussies over mijn onderzoek die wij in je kantoor in Oxford samen hebben gehad waren altijd zo leerzaam en inspirerend. Je stond altijd open voor mijn wellicht wat onconventionele of naïeve ideeën en stimuleerde me om van dit proefschrift echt míjn proefschrift te maken. Ik wil je ook bedanken voor het faciliteren van mijn bezoekjes aan Oxford. Ik voelde me heel welkom en thuis in het Language and Brain Laboratory. Ik heb erg van mijn bezoeken genoten en er veel van geleerd. Niet alleen als wetenschapper, maar ook als mens. Ik dank je ook voor je hulp toen ik tijdens een van mijn periodes in Oxford door een dieptepunt ging. Samen met Henning Reetz heb je mij weer goed op het droge kunnen zetten toen ik bang was kopje onder te gaan. Ik ben jullie beiden bijzonder dankbaar voor jullie tijd en inzet in die periode en voor jullie vertrouwen dat we eruit zouden komen. Het was meer dan jullie hadden hoeven doen.

Tineke, jij had (of maakte) altijd tijd voor mij als het nodig was, ook nadat je niet meer op de Afdeling Nederlands werkte. Je hebt me geholpen met de technische en statistische kanten van dit onderzoek. Jij begreep als een van de weinigen van mijn Nijmeegse collega's wat het inhield om zelf scripts te schrijven voor het draaien van mijn experimenten en mijn analyses. Je hulp hierbij heb ik zeer gewaardeerd en je complimenten voor mijn programmeervaardigheden hebben me goed gedaan. Ik heb ook zelfvertrouwen gehaald uit je positiviteit en je oppeppende woorden die jouw e-mails vaak cadeautjes maakten. Je was voor mij een waardevolle begeleider.

In dit rijtje wil ik ook heel graag Colin Brooks bedanken, de lab-manager van het Language and Brain Laboratory Oxford. Formeel was je geen begeleider, maar in de praktijk eigenlijk wel. Je hebt me op de moeilijkste momenten gemotiveerd om door te zetten om deze marathon te volbrengen. Je stimuleerde me in mezelf te geloven en het mezelf te gunnen dit af te maken.

Je bemoedigende woorden hebben mij ontzettend goed gedaan. Behalve je morele steun heb je me ook technisch en praktisch veel geholpen, zowel tijdens mijn bezoeken aan het lab in Oxford als wanneer ik in Nijmegen was. Zo heb je – zoals het een echte *engineer* betaamt – het voor elkaar gekregen dat ik vanuit mijn eigen kantoor in het Erasmusgebouw op afstand de analysecomputer in Oxford kon gebruiken voor mijn EEG-onderzoek, zodat ik de software tot mijn beschikking had die ik nodig had, maar in Nijmegen niet kon krijgen. Je hebt afgelopen zomer mijn proefschrift uiterst nauwkeurig gelezen en dat deed je ook nog verbazingwekkend snel. Ik dank je voor al je verbetersuggesties en de vanzelfsprekendheid waarmee je dit voor mij wilde doen. Je hebt meer bijgedragen dan je had hoeven doen. Je bent enorm waardevol en ik denk zelfs onmisbaar geweest in dit project.

Mijn dank gaat ook uit naar de leden van de manuscriptcommissie: Marc van Oostendorp, Haike Jacobs, Janet Grijzenhout, Mathias Scharinger en Susanne Brouwer. Bedankt dat jullie mijn proefschrift hebben willen lezen. Ik kijk ernaar uit met jullie van gedachten te wisselen over mijn onderzoek.

Graag wil ik Jeroen Dera, Paul Hulsenboom en Moniek Schaars bedanken voor het beschikbaar stellen van hun stem voor de drie perceptie-experimenten in dit proefschrift. Heel veel dank voor het doorstaan van de spraakopnames. Ik weet dat ik niet snel tevreden was. Moniek, behalve de stem voor mijn kindjesexperiment ben je ook een goede vriendin, die als geen ander begrijpt wat het betekent promotieonderzoek te doen. Bedankt dat je aan mijn zijde wilt staan als paranimf.

Zonder deelnemers geen onderzoeksgegevens. Daarom wil ik de vele volwassenen en de 151 peuters en hun (groot)ouders die aan dit onderzoek hebben deelgenomen hartelijk bedanken. Voor het testen van al deze proefpersonen heb ik hulp gehad van assistenten en stagiaires. In het bijzonder wil ik Elske, Iris, Iris, Lisa en Francie uitdrukkelijk bedanken voor jullie betrokkenheid en al het werk dat jullie erin gestoken hebben. Zonder jullie hulp had dit onderzoek veel langer geduurd.

Enkele collega's van het Language and Brain Laboratory in Oxford hebben mij geholpen bij onder andere het uitvoeren van EEG-metingen en de data-analyse. In het bijzonder wil ik hierbij noemen: Sandra Kotzor, Adam Roberts, Beinan Zhou, Hilary Wynne en Swetlana Schuster. De vanzelfsprekendheid waarmee jullie bijsprongen heb ik zeer gewaardeerd.

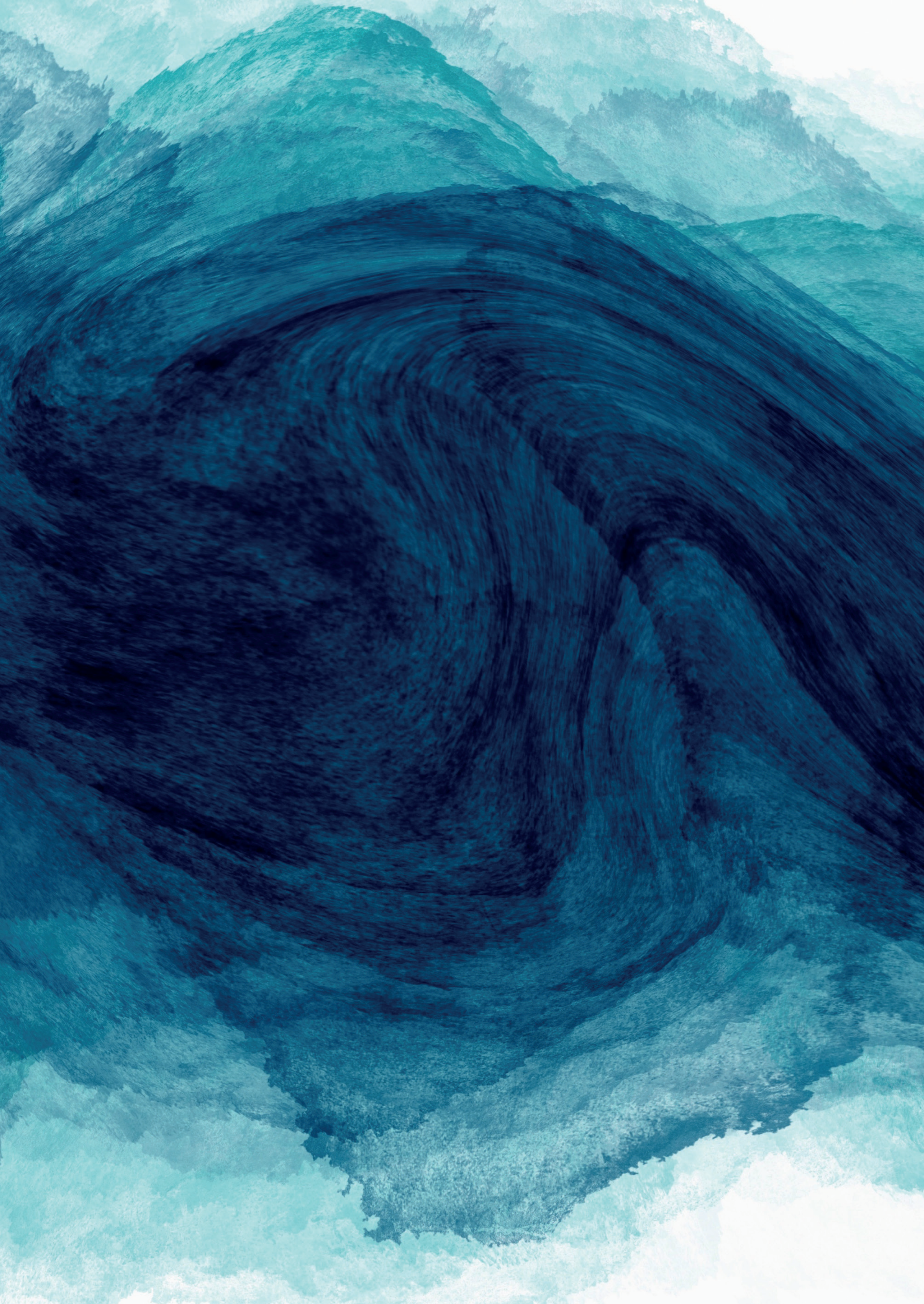
Van mijn collega's in Nijmegen wil ik Susanne Brouwer bedanken voor haar adviezen wat betreft de statistische analyses en Stefanie Ramachers voor het beschikbaar stellen van haar kindjesexperiment voor hergebruik in een iets andere vorm.

Guusje, als kamergenoten hebben wij veel ups en downs met elkaar gedeeld. Toen ik op advies van de bedrijfsarts een ander kantoor kreeg, twijfelde je geen seconde en verhuisde gezellig met me mee. We hebben inmiddels een hechte band opgebouwd, die hopelijk ook nu onze wegen scheiden stand zal houden. Bedankt dat je mijn paranimf wilt zijn.

Mijn lieve familie en vrienden hebben natuurlijk ook bijgedragen door hun steun of juist afleiding tijdens mijn promotietraject. Yvonne, jij was een beschikbaar luisterend oor als het me allemaal even te veel werd. Madelène, jij hielp mij om zoveel mogelijk op mijn eigen kompas te vertrouwen en was het hele traject lang een van mijn grootste fans.

Papa en mama, wat ben ik blij met jullie. Jullie waren erg betrokken bij mijn promotie en wilden er alles over weten. Jullie leefden mee met alle pieken en dalen, adviseerden waar jullie konden en hebben de successen samen met mij gevierd. Een aanzienlijke impact op het verloop van mijn promotie hadden mijn gezondheidsproblemen, die jullie van dichtbij meemaakten. Jullie hebben mij talloze keren in Venlo opgevangen. Ik mocht na elke behandeling bij jullie op de bank liggen tot ik weer (letterlijk) op eigen benen kon staan. Jullie steun was voor mij heel belangrijk om de behandelingen lang genoeg aan te blijven kunnen en daarnaast stug door te gaan met mijn proefschrift op de dagen dat het beter ging. Jullie waren misschien de enigen die echt zagen wat mijn jarenlange therapie inhield en hoe groot de impact was. Gelukkig hebben we ook samen de positieve effecten ervan gezien, die hopelijk nog heel lang mogen duren. Ik kan altijd bij jullie terecht met mijn successen en dieptepunten, waar die ook mee te maken hebben. Dat was tijdens dit promotietraject heel belangrijk.

Lieve Lex, jou heb ik mogen leren kennen in januari 2019. Je was er voor me bij alle hobbels die ik sindsdien ben tegengekomen. Wat ben ik blij dat ik jou aan mijn zijde had tijdens de laatste fase van mijn proefschrift. Ik kon met jou praten over programmeren en zelfs over bomen van fonologische kenmerken en het FUL-model. Je had eindeloos interesse, kon het allemaal volgen en dacht met me mee. Ik was niet altijd het zonnetje in huis, maar dat begreep je en dat maakte jou niet uit. Je probeerde me altijd te troosten of te helpen waar je kon. Elke overwinning, klein of groot, hebben we samen gevierd. Je hebt altijd vertrouwen in mijn kunnen. Je stimuleert me mezelf te zijn - niets meer en niets minder dan mezelf. Dat vind ik heerlijk. En wat lijk je op mij. Je bent het beste wat me ooit is overkomen. Met jou kan ik de wereld aan.



# Biografie

Nadine de Rue is geboren op 12 juli 1990 in Venlo. In 2007 rondde zij het Gymnasium af aan het Valuascollege in Venlo en begon zij aan haar studie Taalwetenschap aan de Radboud Universiteit in Nijmegen. In 2010 behaalde zij deze bachelor cum laude. Tijdens het uitvoeren van haar bacheloronderzoek aan de Sint Maartenskliniek naar de communicatieve ontwikkeling van niet of nauwelijks sprekende meervoudig beperkte kinderen, werd haar onderzoeksinteresse aangewakkerd. Na haar bachelor Taalwetenschap sloot Nadine in 2012 haar master Taal- en Spraakpathologie eveneens cum laude af. Vervolgens voltooide ze in 2013 de tweejarige Research Master Cognitive Neuroscience aan de Donders Graduate School for Cognitive Neuroscience in Nijmegen.

Tijdens haar studie heeft Nadine als studentassistent gewerkt bij het Centre for Language and Speech Technology (CLST) aan de Radboud Universiteit. Na het afsluiten van haar studie heeft zij gewerkt op het Max Planck Instituut (MPI) voor Psycholinguïstiek in Nijmegen op de afdeling Neurobiology of Language, geleid door Peter Hagoort. Ze heeft hier meegewerkt aan allerlei onderzoeken met een verscheidenheid aan methodes en zij kreeg hier de gelegenheid psycholinguïstische experimenten te programmeren. Daarnaast werkte zij als onderzoeksassistent bij de afdeling Gebarentaalwetenschap van de Radboud Universiteit aan het gebarengroeicorpus – een video-corpus (dataset) waarin de verwerving van Nederlandse Gebarentaal als moedertaal door zowel dove als horende kinderen gevolgd wordt.

In oktober 2014 startte Nadine haar promotieonderzoek aan de afdeling Nederlands van de Radboud Universiteit. Zij deed dit onder begeleiding van Prof. dr. Paula Fikkert, dr. Tineke Snijders en Prof. dr. Aditi Lahiri (University of Oxford). Ze heeft voor haar promotieproject een persoonsgebonden beurs (Promoties in de Geesteswetenschappen) toegewezen gekregen van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO). Het onderzoek richt zich op hoe het Nederlandse klinkersysteem in ons brein in elkaar zit. Nadine heeft hierbij onderzoek gedaan bij zowel volwassenen als jonge kinderen en gebruik gemaakt van een verscheidenheid aan onderzoeksmethodes.

Tijdens haar promotietraject heeft Nadine meerdere keren perioden van weken of maanden in Oxford aan haar onderzoek gewerkt in het Language and Brain Laboratory. Ook heeft zij aan de Radboud Universiteit onderwijs gegeven op het gebied van fonetiek en fonologie. Daarnaast heeft Nadine tijdens haar promotietraject met een aparte docentaanstelling bij de afdeling Taalwetenschap van de Radboud Universiteit onderwijs ontwikkeld en lesgegeven op het gebied van Taalontwikkelingsstoornissen.

Sinds mei 2020 werkt Nadine als senior onderzoeker bij Koninklijke Kentalis. Zij doet daar onderzoek met het doel zorg, onderwijs en participatie van dove en slechthorende kinderen en jongeren te verbeteren.