Language-General Biases and Language-Specific Experience Contribute to Phonological Detail in Toddlers' Word Representations

Sho Tsuji
Radboud University Nijmegen and International Max-Planck
Research School for Language Sciences, Nijmegen,
the Netherlands

Paula Fikkert
Radboud University Nijmegen

Naoto Yamane
RIKEN Brain Sciences Institute, Wako, Saitama, Japan

Reiko Mazuka
RIKEN Brain Sciences Institute, Wako, Saitama, Japan, and
Duke University

Although toddlers in their 2nd year of life generally have phonologically detailed representations of words, a consistent lack of sensitivity to certain kinds of phonological changes has been reported. The origin of these insensitivities is poorly understood, and uncovering their cause is crucial for obtaining a complete picture of early phonological development. The present study explored the origins of the insensitivity to the change from coronal to labial consonants. In cross-linguistic research, we assessed to what extent this insensitivity is language-specific (or would show both in learners of Dutch and a very different language like Japanese), and contrast/direction-specific to the coronal-to-labial change (or would also extend to the coronal-to-dorsal change). We measured Dutch and Japanese 18-month-old toddlers' sensitivity to labial and dorsal mispronunciations of newly learned coronal-initial words. Both Dutch and Japanese toddlers showed reduced sensitivity to the coronal-to-labial change, although this effect was more pronounced in Dutch toddlers. The lack of sensitivity was also specific to the coronal-to-labial change because toddlers from both language backgrounds were highly sensitive to dorsal mispronunciations. Combined with results from previous studies, the present outcomes are most consistent with an early, language-general bias specific to the coronal-to-labial change, which is modified by the properties of toddlers' early, language-specific lexicon.

Keywords: word learning, phonological development, perceptual asymmetries, mispronunciation detection

Infants become highly proficient in discriminating native speech sound categories during their first year of life. This ability is, however, not always straightforwardly applied in lexical tasks. For instance, English-learning toddlers sometimes fail to notice that an object previously labeled [bɪ] was subsequently labeled [dɪ] in a word-learning task (Stager & Werker, 1997). The reason for this discrepancy has been a matter of much debate, and possible explanations have focused on the difference between phonetic and phonological representations (the former allowing discrimination of a contrast, the latter being required to apply this ability in lexical contexts) or on differences in task demands between discrimination tasks and lexical tasks.

More recent studies add a new perspective to this debate by suggesting that results such as in Stager and Werker (1997) are related to toddlers' insensitivity to specific, directional changes: Multiple studies have found an asymmetry in toddlers' perception, such that they have no difficulties detecting a change from a word including a labial speech sound to a word containing a coronal speech sound (labial-to-coronal change), but do not succeed in detecting the opposite coronal-to-labial change (Altvater-Mackensen, van der Feest, & Fikkert, 2014; Fennel, van der Feest, & Spring, 2011; Fikkert, 2010; van der Feest & Fikkert, in press).

Despite these multiple findings (which are discussed in more detail below), the cause and scope of the perceptual insensitivity to the coronal-to-labial change are unclear: Is it a phenomenon caused by exposure to specific languages or is it more language general? Is it specific to the coronal-to-labial change or is it also occurring for other changes, for instance coronal-to-dorsal? Be-

cause toddlers, by and large, have been found to be highly sensitive to phonological detail in lexical tasks, understanding the causes of this consistent insensitivity is crucial to gain a complete picture of early phonological and lexical development. The present study investigates the origins of the reported perceptual insensitivity by comparing the perception of toddlers from different language backgrounds (Dutch, Japanese) and by testing sensitivity not only to the coronal-to-labial but also to the coronal-to-dorsal change.

In the following, we will first summarize what is known about toddlers' sensitivity to phonological detail, demonstrating that, in general, they are highly sensitive to differences in native speech sounds. We will proceed to describe in detail the findings on perceptual insensitivities and possible causes suggested, followed by the present study.

Phonological Detail in Early Lexical Representations

Numerous studies have demonstrated that toddlers have detailed phonological representations of familiar or newly learned words. This matter has often been investigated with the preferential looking procedure (Swingley & Aslin, 2000), a paradigm in which two pictures of known objects are presented side by side on a screen while one of them is named. Toddlers are expected to look to the named object in this correct pronunciation (CP) condition. Provided toddlers are sensitive to phonological detail, their recognition should be hindered when presented with a mispronounced version of the object name, leading to a lower amount of object looks compared with the CP condition. Using this or related paradigms, it has been demonstrated that toddlers as young as 11 months of age are sensitive to phonological detail in consonants (e.g., Bailey & Plunkett, 2002; Swingley, 2005). They can access such detail not only at word onset but also at word-medial or word-final position (Nazzi, 2005; Nazzi & Bertoncini, 2009; Swingley, 2003; Swingley, 2009). Whether toddlers are equally sensitive to voice, manner and place of articulation mispronunciations (MPs; cf. Mani & Plunkett, 2010a; White & Morgan, 2008), whether their sensitivity to MPs of one or more phonological features is graded (cf. Bailey & Plunkett, 2002; White & Morgan, 2008), and whether their sensitivity to phonological detail in vowels is comparable to that in consonants (cf. Mani & Plunkett, 2010a; Nazzi, 2005) is a matter of debate and reviewed in detail elsewhere (e.g., Altvater-Mackensen et al., 2014). In any event, the above summary documents toddlers' robust sensitivity to consonantal detail in lexical tasks.

Asymmetries in Early Lexical Representations

In contrast to the above findings, a word learning study with English-learning 14-month-old toddlers found no such sensitivity to phonological detail (Stager & Werker, 1997). Toddlers were first habituated either to two objects (Experiment 1; one labeled [bɪ] and another labeled [dɪ]) or to one object (Experiment 2; labeled either [bɪ] or [dɪ]), and subsequently tested on their ability to recognize whether an object label had changed such that either the object previously associated with [bɪ] was now presented together with the label [dɪ] or vice versa. Whereas the original study did not report directional effects, a recent reanalysis of a series of related experiments in Fennell and Waxman (2010)

suggested that toddlers are insensitive to the change in one particular direction. Toddlers for whom the object label starting with [b] changed to starting with [d] were able to perceive the change, but not toddlers for whom [d] changed to [b] (Fennel, van der Feest, & Spring, 2011). This reanalysis was additionally backed by Fikkert's (2010) modified replication of the Stager and Werker (1997) study with Dutch 17-month-old toddlers. After being habituated to an object labeled [bɪn] or [dɪn], they were presented with the same object paired with the respectively different label. Toddlers habituated to [bɪn] detected the change to [dɪn], but toddlers habituated to [dɪn] did not show evidence of detecting the change to [bɪn].

Further corroborating evidence comes from a series of studies on word recognition in Dutch toddlers. Assessed in a preferential looking paradigm, 20- and 24-month-old toddlers were sensitive to the change from labial-initial/pus/(*cat*) to its coronal MP [tus] but not to the change from coronal-initial/tɑnt/(*tooth*) to its labial MP [pɑnt] (van der Feest & Fikkert, in press). This pattern was replicated with 18-month-old toddlers, who were able to detect the change from labial-initial/vɪs/(*fish*) to its coronal MP [zɪs] but not the change from coronal-initial/ze:p/(*soap*) to its labial MP [ve:p] (Altvater-Mackensen et al., 2014).

In summary, a labial-coronal perceptual asymmetry has been documented in different tasks, with different stimuli, and in toddlers up to the age of 24 months. This replication of the asymmetry across paradigms, across stimuli, and up to a relatively high age is difficult to reconcile with an explanation purely involving age or task demand (for a similar discussion, cf. Werker, Fennell, Corcoran, & Stager, 2002; Yoshida, Fennell, Swingley, & Werker, 2009). Toddlers' ability to access the necessary phonological detail in one direction of change also speaks to their general ability to perform the tasks at hand. Therefore, these results suggest a genuine lack of sensitivity to the coronal-to-labial change in lexical tasks for Dutch- and English-learning toddlers (but cf. Ren & Morgan, 2012, for a failure to find such an asymmetry in English-learning toddlers).

Possible Causes of the Labial-Coronal Asymmetry

Several, not necessarily mutually exclusive accounts, have been brought forward to explain the origins of the labial-coronal perceptual asymmetry. The accounts range from the assumption of abstract phonological representations to early universal biases to accounts that consider toddlers' language input and their own productions. The most prominent account suggests that the asymmetry is tied to the special status of coronals in the phonologies of the world (Paradis & Prunet, 1991). The featurally underspecified lexicon (Lahiri & Reetz, 2010) accounts for the asymmetry by assuming sparse and abstract lexical representations in which not all phonological features are specified. Coronal place of articulation is assumed to be the default, unmarked place and therefore to be underspecified in the mental lexicon. This assumption predicts the documented perceptual asymmetries such that labial MPs like [pɑnt] are accepted as an instance of coronal/tɑnt/, but coronal MPs like [tus] would not be accepted as an instance of labial/pus/. A perceived labial stop is not mismatching with the underspecified representation of a coronal stop, hence [pɑnt] for/tɑnt/is acceptable, whereas a perceived coronal mismatches with a specified labial, hence [tus] for/pus/is not accepted. Under the assumption that coronals universally are the default place of articulation, the

underspecification account would further predict the same asymmetry for the coronal-dorsal contrast, and for toddlers across language backgrounds [cf. prediction (a) in Table 1]. On a slightly different account, Fikkert (2010) suggested that toddlers start out with sparsely specified lexical phonological representations and gradually add detail in the course of development.

This account's prediction for the coronal-to-labial change in Dutch conforms to findings of previous studies, but the extent to which the perceptual insensitivities generalize to the coronal-to-dorsal change and to Japanese toddlers remains to be tested.

Although the featurally underspecified lexicon approach crucially assumes lexical representations, recent findings suggest that language-general perceptual biases could play a role in the above-described asymmetries already at a prelexical stage of development. A cross-linguistic discrimination study documented that both 4- to 6-month-old Dutch and Japanese infants were able to detect a change from labial [ɔmpa] to coronal [ɔnta], but not from coronal [ɔnta] to labial [ɔmpa] (Tsuji, Mazuka, Cristia, & Fikkert, 2015; see also Dijkstra & Fikkert, 2011, for similar evidence using the word-initial contrast [pa:n-ta:n]). This similarity in discrimination performance of infants learning two unrelated languages at an age at which they have not yet acquired language-specific consonant representations (cf. Kuhl, 2004) is suggestive of a language-general perceptual bias as the origin for the lack of sensitivity to the coronal-to-labial change documented in Dutch toddlers. If this bias persists, we would expect Japanese toddlers to also continue showing the same insensitivity later on [cf. prediction (b) in Table 1; note that no data on infants' perception of the coronal-to-dorsal change exist].

A third account takes toddlers' language input into account, in particular the high frequency of coronals in Dutch and related languages. Recent evidence suggests that discrimination in the direction from more frequent to less frequent phonemes is harder than vice versa for infants by the age of 12 months (Pons, Albareda-Castellot, & Sebastián-Gallés, 2012), potentially because more frequent native phonemes are treated as referents, making discrimination toward less frequent phonemes more difficult than the opposite. In Dutch, this would predict discrimination from coronal to labial or dorsal place to be more difficult than the opposite, because coronal plosives are more frequent than labial or dorsal plosives in infant-directed speech under most counts (cf.

Table 2; counts derived from van de Weijer corpus, van de Weijer, 1998). The predictions for Japanese differ, because coronal plosives are more frequent than labial but less frequent than dorsal plosives under most counts (cf. Table 2; counts derived from R-JMICC; Mazuka, Igarashi, & Nishikawa, 2006). Consequently, discrimination of the coronal-to-labial change is predicted to be difficult but not discrimination of the coronal-to-dorsal change.

Frequency-based predictions need to be treated with caution, however, as it is unclear what kind of frequency count—overall frequency or subsets containing word-initial phonemes or content words—would matter. We calculated token frequencies for all these types of count and based our predictions on the dominant pattern in the respective language [cf. prediction (c) in Table 1]. It is, however, possible that the frequencies in word-initial positions of content words, which are especially salient to toddlers, could be more relevant than our "majority count" (de Boysson-Bardies & Vihman, 1991; Shi & Werker, 2001; Vihman, Kay, de Boysson-Bardies, Durand, & Sundberg, 1994). Furthermore, for toddlers in the middle of vocabulary acquisition, frequencies of content word types in their input, as well as in their own receptive and productive inventories, might be important factors contributing to the structure of their phonological representations.

A fourth prediction is based on toddlers' own productions. Recent studies suggest a relationship between toddlers' early productions and their perception, such that they pay less attention to phonemes they produce frequently (e.g., DePaolis, Vihman, & Keren-Portney, 2011; DePaolis, Vihman, & Nakai, 2013). With regard to MPs, this could mean that toddlers pay less attention to MPs if they consist of a frequently produced phoneme. Consequently, they might accept them more readily than MPs that consist of a less frequently produced phoneme.

Indeed, it is known that Dutch toddlers predominantly produce labial-initial words early on (Fikkert & Levelt, 2008), but Japanese toddlers' early productions contain a high number of dorsals (de Boysson-Bardies & Vihman, 1991). This pattern can also be observed in vocabulary questionnaires in both languages (for Dutch, Schlichting & Spelberg, 2002; for Japanese, Japanese Mac-Arthur CDI, Ogura & Watamaki, 2004): In plosive-initial words produced by at least 20% of 17- or 19-month-old Dutch toddlers, labial-initial words are most frequent (17 months: 15; 19 months: 32), followed by coronal-initial (17 months: 7; 19 months: 19), and dorsal-initial (17 months: 4; 19 months: 13) words. In contrast, the count for 18-month-old Japanese toddlers reveals that dorsal-initial words are most frequent (11), followed closely by labial-initial (10) and lastly coronal-initial words (6). Research on speech errors also shows comparable differences: [1] Dutch toddlers are reported to predominantly make "fronting" errors, which often means replacing a coronal or dorsal consonant with a labial (Fikkert & Levelt, 2008). In contrast, Japanese toddlers have been found to make "backing" errors, replacing coronal with more back conso-

Table 1

Predictions on the Detection of Labial and Dorsal Mispronunciations of Coronal-Initial Words Made by the Four Different Accounts for Dutch and Japanese

Predictions	Dutch		Japanese	
	Coronal to labial	Coronal to dorsal	Coronal to labial	Coronal to dorsal
(a) Underspecification	no	no	no	no
(b) Early bias	no	?	no	?
(c) Frequency	no	no	no	yes
(d) Production	no	yes	yes	no
Experimental results	no	yes	yes	yes

Note. The last line summarizes the results from Experiments 1 and 2. "No" refers to a lack of sensitivity to the contrast, "yes" refers to contrast detection, and "?" refers to cases where no predictions can be made.

---

[1] Note, however, that the results on speech errors might partly reflect language-specific differences in adult listeners' perception of toddlers' ambiguous productions as suggested by the findings in Li, Munson, Edwards, Yoneyama, and Hall (2011). Thus, adult native speakers of Dutch may perceive children's productions that are intermediate between coronal and labial as labial, whereas adult native speakers of Japanese may perceive children's productions that are intermediate between coronal and dorsal as dorsal.

Table 2
Frequency of Plosives by Place of Articulation

Type of count	Labial (%)	Coronal (%)	Dorsal (%)
Dutch			
Token			
AW, all	20.6	**52.4**	27.0
AW, ini	27.6	**43.5**	28.9
CW, all	18.1	**48.0**	33.8
CW, ini	28.0	31.7	**40.3**
Type			
CW, ini	**47.9**	26.5	25.6
Japanese			
Token			
AW, all	13.8	42.7	**43.5**
AW, ini	11.0	**51.4**	37.6
CW, all	21.4	30.2	**48.3**
CW, ini	24.6	29.9	**45.5**
Type			
CW, ini	30.4	24.6	**45.0**

Note. Dutch counts derived from van der Weijer Corpus (van de Weijer, 1998), Japanese counts from R-JMICC (Mazuka, Igarashi, & Nishikawa, 2006). AW = all words; CW = content words; all: phonemes in all positions of the word; ini: phonemes in word-initial position.

nants (Edwards & Beckman, 2008). Thus, overall Dutch toddlers know more words starting with labial-initial plosives, whereas Japanese toddlers know more words starting with dorsal-initial plosives. Although this pattern does not conform to the input token frequencies reported above, type frequencies of content words (which might better reflect the words toddlers know and produce) show that word-initial plosives are most frequently labials in Dutch but dorsals in Japanese (cf. Table 2). The production account would, thus, predict that labial MPs are accepted as instances of coronal-initial words by Dutch but not by Japanese toddlers. The reverse would hold for dorsal MPs, which should be accepted as instances of coronal-initial words by Japanese but not by Dutch toddlers [cf. prediction (d) in Table 1].

The Current Study

Using a word-learning task, the current study investigated Dutch and Japanese toddlers' sensitivity to phonological detail in coronal-to-labial and coronal-to-dorsal changes. Toddlers first learned two novel coronal-initial word–object associations and were subsequently tested on their sensitivity to labial and dorsal

MPs of the target names in a preferential looking paradigm. Experiment 1 assessed Dutch toddlers, and Experiment 2 assessed Japanese toddlers. Toddlers' early vocabularies were measured by asking Dutch caregivers to fill in the Dutch Communicative Development Inventory N-CDI (short version; Zink & Lejaegere, 2003) and Japanese caregivers to fill in the Japanese MacArthur Communicative Development Inventory (Ogura & Watamaki, 2004).

Experiment 1: Dutch Toddlers

Participants

Thirty-one monolingual Dutch toddlers were included in the final sample (mean age = 18.77 months, range = 18.31–19.13 months, 15 female). They were recruited and tested in the Netherlands. Caregivers signed an informed consent and received a picture book or a small monetary compensation for their participation. Twenty-three additional toddlers were excluded due to fussiness and not completing the experiment (7), technical problems (8), or due to the exclusion criteria explained below (8).

Stimuli

Target stimuli consisted of two word-object pairings and their respective MPs. Two stuffed animals were used as target stimulus objects (cf. Figure 1). They had distinct colors and shapes to make them easily discriminable.

The names associated with the two stuffed animals were coronal-initial [ta:səl] and [tɑno:]. Low unrounded vowels were chosen to follow the critical word-initial coronal consonant, and the place of articulation of the two word-medial consonants was matched. Labial MPs were [pa:səl] and [pɑno:], and dorsal MPs were [ka:sel] and [kɑno:]. Stimuli were recorded by a native female speaker of Dutch in a child-directed register, and all tokens were embedded in carrier phrases. The auditory target stimuli were cross-spliced into the carrier phrases such that CPs, labial MPs, and dorsal MPs were embedded into identical carrier phrases.

The experiment consisted of a learning and a test phase, with the learning phase being subdivided into a live and screen learning phase (cf. Figure 1). Live learning provides a more naturalistic situation that possibly has an advantage over screen learning (DeLoache et al., 2010), and a subsequent screen learning phase made toddlers familiar with seeing the target objects on screen.
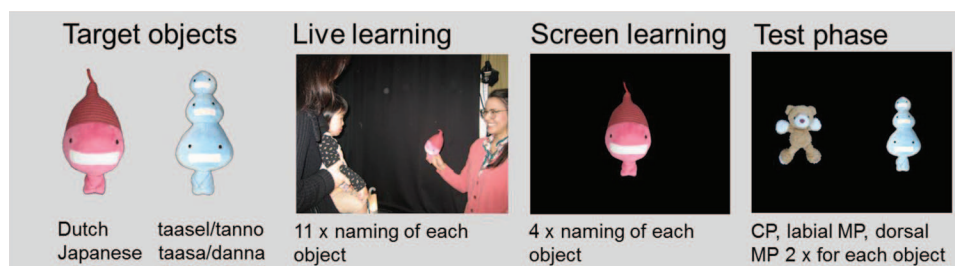

Figure 1. Target objects and procedure for Experiments 1 and 2. The authors received signed consent for the child's and caregiver's likeness to be published in this article. CP = correct pronunciations; MP = mispronunciations. See the online article for the color version of this figure.

Whether the blue or the pink object was named [taːsel] or [tɑnnoː], and which object was presented first (pink or blue) was counterbalanced between participants.

During live learning, the actual stuffed animals served as visual stimuli. The experimenter, a female native speaker of Dutch, named each object 11 times in scripted carrier phrases such as "Do you want to play with the [target]?" Subsequently, toddlers were exposed to three screen learning trials per object. In each trial, the photograph of one of the target objects was presented centrally against black background for 4 s while slightly increasing and decreasing in size. Each object was named twice in the first trial and once in each of the remaining two, resulting in four naming instances during 12 s of exposure per object. A different auditory token was used for each naming instance.

In test trials, the photograph of one of the objects was presented side by side with the photograph of an unrelated stuffed animal against a black background for 6.5 s. The pair slowly moved up and down while one of the objects was named. The onset of the initial consonant of the target name was always at 3.7 s. As it was impossible to cross-splice the same instance of the definite article [də] (which always preceded the target name) onto the labial-, coronal-, or dorsal-initial target tokens due to coarticulation, target names were cross-spliced into the carrier phrases together with their article (e.g., "Can you see [the target]?"). Different auditory tokens were used in each trial, with auditory characteristics matched between CPs and MPs (cf. Table 3). Each of the three pronunciation conditions occurred twice for each of the two target names, resulting in a total of 12 test trials. The pink object was always paired with an elephant, and the blue object with a bear with target side counterbalanced. In addition, there were 12 filler

trials. The elephant and the bear were named twice each. The remaining eight filler trials consisted of two additional pairs of animals (cat-giraffe, dog-pig; each named twice with side of presentation counterbalanced). Four pseudorandom trial orders were created and counterbalanced between toddlers.

Procedure

Toddlers were tested in a sound-attenuated room while sitting on their caregivers' lap. The experimenter removed one of the target objects from behind a curtain, showing and naming it according to the script. The child was allowed to touch and hold the object. The caregiver was instructed to remain silent but was encouraged to smile and look at child and object. In case the child was afraid of the object, the caregiver was also encouraged to hold it for a little while before the experimenter initiated the learning phase. Immediately following the live learning phase, the child and caregiver were seated in front of a Tobii T60 eyetracker. The experimenter monitored and initiated trials from a separate computer screen in the same room. Caregiver and experimenter listened to masking music throughout the on-screen experimental phase. After calibration, the six screen learning trials and the 24 experimental trials were presented. Each trial was preceded by a smiley face in the screen center and was initiated once the child fixated the smiley face. A short movie of a duck was presented after the learning trials and after the first half of the experimental trials.

Data Preprocessing

Exclusion criteria on trial and participant level were applied. To exclude trials in which toddlers only spuriously looked, test trials in which they looked less than 500 ms of the 2,000 ms following target word onset (367–2,367 ms) to anywhere on screen were excluded. This excluded 18% of test trials. To exclude toddlers that were not attentive during screen learning, we excluded five toddlers that had accumulated less than 4 s (of 12 s) of looking time to either of the two objects. Additionally, we excluded three toddlers who did not contribute at least one valid trial per condition during the test phase.[2]

Data Analysis

Studies using the preferential looking paradigm have mostly analyzed outcome data by averaging over a given time window. In the present study, however, the trajectory of the naming effect in the CP condition differed between Dutch toddlers in Experiment 1 (peak around 850 ms after target word onset) and Japanese toddlers in Experiment 2 (peak around 1,200 ms). To make the critical comparison between the MP conditions and the CP condition in a comparable way for the two language groups, we chose a common time window based on the pooled data of both language groups as a first step. We pooled the data from the two experiments in the

Table 3
Acoustic Measurements of Experimental Stimuli

Stimulus type	Length (ms)	Pitch (Hz)
Dutch		
taasel		
Screen Learning	729	306
CP	724	335
Labial MP	728	343
Dorsal MP	712	350
tanno		
Screen Learning	714	327
CP	707	324
Labial MP	691	329
Dorsal MP	674	321
Japanese		
taasa		
Screen Learning	298	347
CP	309	388
Labial MP	317	368
Dorsal MP	297	360
danna		
Screen Learning	315	342
CP	349	349
Labial MP	386	353
Dorsal MP	343	356

Note. Numbers in learning rows are the means over all four learning stimuli, numbers in all other rows are means over the two respective test stimuli. Length is measured over the whole token, and pitch refers to the mean pitch over vowels. CP = correct pronunciations; MP = mispronunciations.

---

[2] We used rather lenient exclusion criteria to include as many participants as possible. If we instead included toddlers only if they looked to the screen at least 1,000 ms (50%) of the postnaming time window and contributed at least 50% of valid trials overall and in each condition during the test phase, we would have excluded 7 toddlers and 23% of their trials.

367–2,367 ms following word onset and averaged the looks to target in the CP condition, which can be considered the baseline condition, over 100 ms time slices. We then determined the time slice with maximum target looks and defined the analysis time window as the 1,000 ms around the peak time slice. This procedure resulted in an analysis window of 500 to 1,500 ms after word onset, which was applied to both data sets. We analyzed the data of this time window with growth curve analysis (GCA, Mirman, 2014). GCA accounts for the dynamic nature of gaze data by not only assessing overall differences in looking times but additionally differences in the shape and latency of the gaze curve. GCA was conducted with the lme4 package (Bates, Maechler, & Bolker, 2012) in R (R Core Team, 2012). The time course of the naming effect was captured with first (linear) and second order (quadratic) orthogonal polynomials and with fixed effects of condition (CP, labial MP, dorsal MP) on all time terms. To account for individual differences, we allowed random effects of participant and participant-by-condition on all time terms. The log-odd transformed ratio of looks to target were the dependent variable. The CP condition was defined as the baseline against which the other conditions were compared.

One caveat of the above analysis is that our choice of matching time windows across experiments might have reduced sensitivity for one or both language groups. As a second step, we therefore applied the procedure for finding a 1,000 ms time window of analysis separately to the data of each language group and reanalyzed the data based on these language-specific time-windows (300–1,300 ms for the Dutch and 800–1,800 ms for the Japanese sample).

Having to base our analyses on two different time windows is because there is no objectively established criterion for the choice of an appropriate time window, and other authors might have made different choices. We therefore conducted a complementary analysis adopting a nonparametric statistical test (NPST), which determines the time stretches in which conditions differ from each other in a bottom-up way. This procedure was initially introduced for the analysis of event-related potential (ERP) data (Maris & Oostenveld, 2007) and has been applied to child eye-tracking data recently (Von Holzen & Mani, 2012). T tests on the difference between two conditions were calculated for each time point in the 2 s following word onset (367–2,367 ms). On the basis of these t tests, time-adjacent clusters of time points whose t values exceeded a certain threshold ($p < .05$, two-tailed) and had the same sign were identified. To account for the multiple comparison problem, Monte-Carlo resampling was then applied, and the trials of the original data set were randomly reassigned to the three conditions 1,000 times. For each of these resampled data sets, the largest cluster as determined by the sum of its t values was identified. Finally, we tested whether the largest cluster of the original sample was significantly different from chance by comparing it to the largest clusters of the 1,000 resampled data sets. This procedure identifies time intervals in which two conditions differ from each other while controlling for multiple comparisons (see Maris & Oostenveld, 2007, for details).

The above analyses are indicative of the difference between conditions but do not assess whether or not there was a naming effect (an increased amount of target looks after hearing the target word) for the CP and the two MPs. Therefore, we supplemented these analyses by comparing the difference scores between pro-

portion of target looks in the 3,700-ms window before naming and the proportion of target looks in the common and language-specific postnaming time window to chance level.

Because there exists no clear consensus for effect-size calculation for both of these analysis methods, we report standard Cohen's d effect sizes for the common time window of 500–1,500 ms.

Results and Discussion

The top part of Figure 2 shows the time course of looks to target for the CP and MP conditions for Dutch toddlers. The averaged percentage of looks to target during the common time window chosen for GCA (see Figure 2) was 62.0% in the CP condition, 58.7% in the labial MP condition, and 53.6% in the dorsal MP condition. The labial MP condition neither had an effect on the intercept ($b = -0.21$, $t = -1.23$. $p = .217$) nor on the quadratic time term ($b = -0.051$, $t = -0.21$, $p = .808$), indicating no difference to the CP condition. By contrast, the dorsal MP condition had a significant effect on the intercept ($b = -0.367$, $t = -2.16$, $p = .031$) but not on the quadratic polynomial ($b = -0.121$, $t = -0.57$, $p = .567$). Cohen's d indicated an effect size of $d = 0.262$ for the difference between CP and labial MP, and $d = 0.439$ for the difference between CP and dorsal MP.

In the language-specific Dutch time window, the averaged percentage of looks to target was 62.2% in the CP condition, 58.3% in the labial MP condition, and 52.5% in the dorsal MP condition. Consistent with the analysis based on the common time window, the labial MP condition neither had an effect on the intercept ($b = -0.29$, $t = -1.57$. $p = .117$) nor on the quadratic time term ($b = -0.069$, $t = -0.38$, $p = .703$). The dorsal MP condition again had a significant effect on the intercept ($b = -0.362$, $t = -2.01$, $p = .045$) but not on the quadratic polynomial ($b = -0.147$, $t = -0.82$, $p = .415$). The effect size was $d = 0.272$ for the difference between CP and labial MP, and $d = 0.396$ for the difference between CP and dorsal MP. The main effect of the dorsal MP in both analyses reflects that toddlers looked less to the target object after hearing the dorsal MP compared with the CP.[3]

The analysis on the difference score between prenaming and postnaming window was consistent with the GCA analyses, showing a significant increase in target looks for the CP in both the common and language-specific time windows: common, $m = 0.12$, $t(29) = 2.89$, $p = .007$; language-specific, $m = 0.12$, $t(30) = 3.92$, $p = .003$. A naming effect was also observed for the labial MP: common, $m = 0.10$, $t(29) = 2.21$, $p = .035$; language-specific, $m = 0.09$, $t(30) = 1.87$, $p = .071$. However, it was not observed for the dorsal MP: common, $m = 0.02$, $t(29) = 0.42$, $p = .676$; language-specific, $m = 0.02$, $t(30) = 0.34$, $p = .740$.

The NPST also backed the results of the above analyses, revealing no significant differences between toddlers' looks to target in the labial MP condition compared to the CP condition in the GCA time-window. In a later time window between 1,651 and 1,768 ms, however, their target looks in the labial MP condition (62.7%) increased, marginally exceeding the CP condition (49.2%), $t = -24.86$, $p = .077$. By contrast, toddlers showed a clear MP effect for the dorsal MP condition in the GCA time window, with significantly

---

[3] Repeating these analyses with more stringent exclusion criteria (cf. footnote 2) lead to qualitatively the same results.
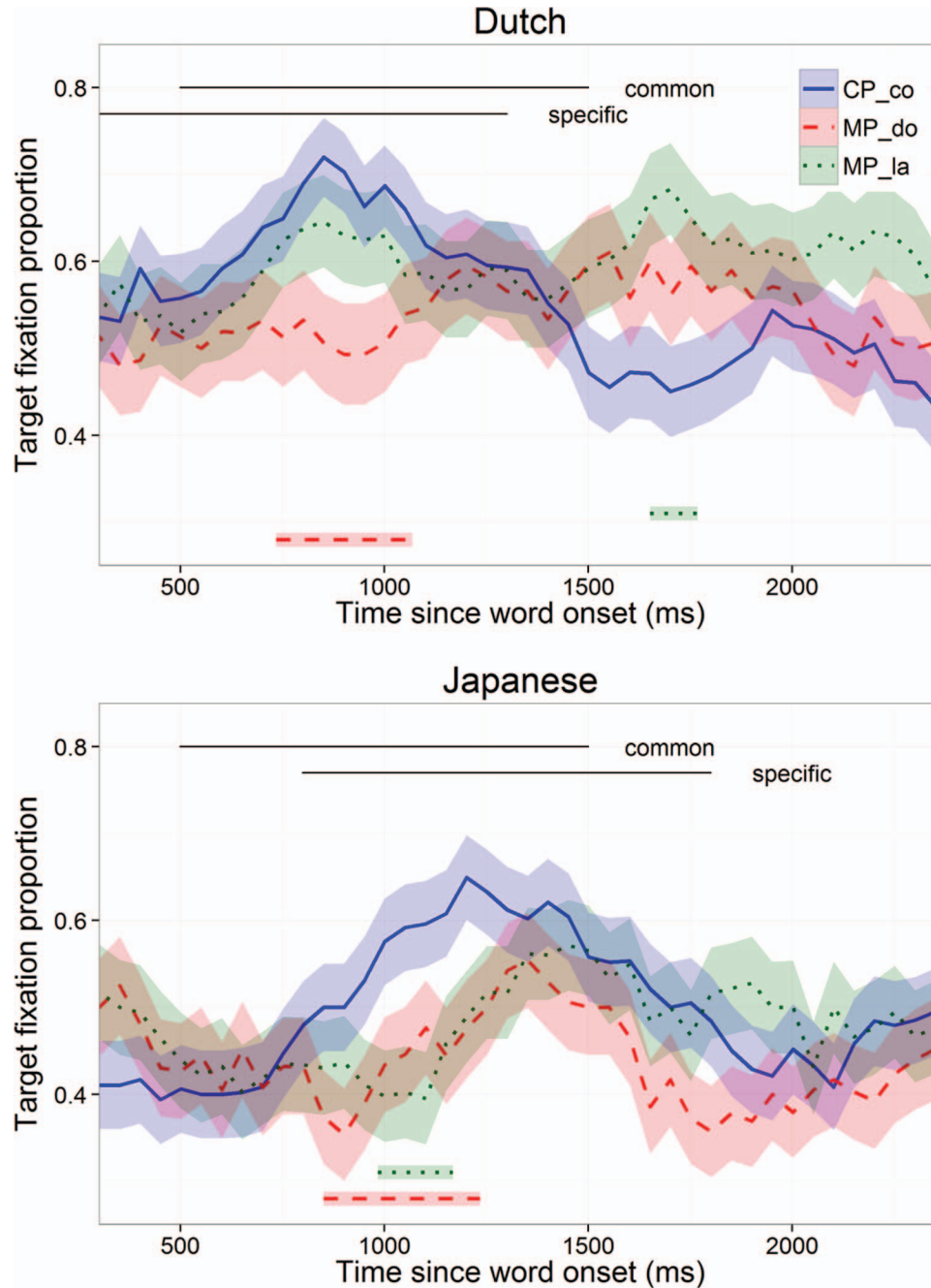
Figure 2. Time course of gaze after word onset for correct pronunciation (CP), dorsal mispronunciation (MP dorsal), and labial mispronunciation (MP labial), separately for Dutch (above) and Japanese (below) toddlers. The black solid lines indicate the time windows (common or language specific) for the growth-curve analysis. The dotted lines at the bottom of the graphs signify the time intervals that resulted in significant differences in the nonparametric test. Shaded areas indicate $\pm SE$. See the online article for the color version of this figure.

fewer looks to target in the MP (51.1%) compared to the CP (68.4%) condition between 734 and 1,068 ms, $t = 67.70$, $p = .004$.

For the analysis of the N-CDI, we counted the number of plosive-initial words for each place of articulation in each child's receptive and productive vocabularies and calculated the mean number of words per place (see Table 4). The plosive-initial words

toddlers understood were mostly dorsal, followed by labial and coronal words. Words produced were most frequently labial, followed by dorsal and coronal, reflecting what has been found in studies on early production (cf. Fikkert & Levelt, 2008).

The converging results on the labial MP condition are in line with previous findings: Dutch toddlers did not look less to the

Table 4

Mean Number of Plosive-Initial Words by Place of Articulation in Children's Early Inventories According to Parental Report

Language	Words	Labial	Coronal	Dorsal
Dutch	understood	5.15	4.15	**5.89**
	produced	**3.44**	2.00	3.30
Japanese	understood	7.72	13.45	**20.79**
	produced	5.00	3.83	**7.90**

Note. Dutch = Zink & Lejaegere, 2003; Japanese = Ogura & Watamaki, 2004. Boldface type indicates highest number in a row. Absolute frequencies are because a full version of the questionnaire was used in Japanese but not in Dutch.

target object when it was mispronounced than when it was correctly pronounced, thus accepting labial MPs as instances of previously learned coronal-initial words. The NPST did, however, reveal a response pattern not reported in previous studies: Toddlers' gaze trajectory to the labial MPs was not identical to the trajectory toward the CPs, but rather suggested delayed "recognition": Although looks to target in the CP condition peaked early and then returned to chance, looks to target in the labial MP condition showed a tendency to increase later on. This finding suggests that toddlers were not absolutely insensitive to the difference between labial- and coronal-initial stimuli, even though they accepted labial MPs as instances of coronal-initial words. Previous studies might have failed to detect such differences between the two conditions, either because they analyzed time-windows centered around the peak response as we did for GCA or because they reported averaged responses over a larger time window (which, for the current data set in the time window of 367–2,367 ms after target-word onset, would also lead to very similar percentages: 53.1% for CP, and 53.5% for labial MP).

Although the results for the coronal-to-labial change are consistent with predictions from all accounts in Table 1, toddlers' clear sensitivity to dorsal MPs that was confirmed in both analyses is inconsistent with the FUL and the frequency account: Both would have predicted Dutch toddlers to accept dorsal MPs as instances of coronal-initial words. The results of Experiment 1 therefore exclude both as possible explanations for toddlers' differential sensitivity toward place of articulation changes. The outcomes are, however, compatible with both the early language-general bias (which has only been attested for the coronal-to-labial change) and the production account. These accounts would, however, make diverging predictions for Japanese toddlers: If the lack of sensitivity to the coronal-to-labial change was indeed exclusively reflecting an early bias, we would expect Japanese toddlers' perceptual patterns to resemble those of Dutch toddlers. An influence of early word knowledge would, however, predict Japanese toddlers to be sensitive to the coronal-to-labial (but not the coronal-to-dorsal) change. To disentangle these possibilities, Experiment 2 assessed Japanese toddlers' perceptual patterns.

Experiment 2: Japanese Toddlers

Participants

Twenty-nine monolingual Japanese toddlers were included in the final analysis (mean age = 18.46 months, range = 18.02– 19.04 months, 13 female). They were recruited and tested in the Tokyo area of Japan. Caregivers signed an informed consent, and received a book voucher in return for their participation. Seven additional toddlers were tested but not included into analysis due to fussiness and not completing the experiment (1), poor tracking of gaze (1), equipment error (1), experimenter failure (1) or due to the exclusion criteria detailed out below (3).

Stimuli

The Japanese object names were [ta:sa] and [daNna] for CPs and [pa:sa]/[baNna] and [ka:sa]/[gaNna] for the respective labial and dorsal MPs. These nonwords were matched as closely as possible to their Dutch counterparts while making them sound natural in Japanese. Auditory stimuli were recorded by a female native speaker of Japanese in child-directed register. Visual stimuli were the same as in Dutch.

Procedure

Japanese toddlers were tested on a Tobii 60 XL, which has a larger screen than the eye tracker used in Experiment 1. To make the positioning and size of stimuli identical to Experiment 1, a subset of the screen was used that was identical in size to the screen used in Experiment 1. The experimental procedure was identical to Experiment 1.

Data Preprocessing

The same exclusion criteria as in Experiment 1 were applied. Exclusion of trials with less than 25% of looks to target after naming resulted in the exclusion of 14% of trials. Two participants were excluded who had accumulated less than 4 s of looking time to each of the toys in the screen learning phase, and one participant was excluded for having less than one trial per condition in the test phase.[4]

Data Analysis

Analyses were exactly the same as in Experiment 1.

Results and Discussion

The bottom part of Figure 2 shows the time course of looks to target for Japanese toddlers. The averaged percentage of looks to target during the common GCA time window was 53.7% in the CP condition, 46.0% in the labial MP condition, and 45.5% in the dorsal MP condition. The GCA showed a significant effect of the labial MP on the intercept ($b = -0.403$, $t = -2.57$, $p = .010$) but no significant effect on the quadratic time term ($b = -0.195$, $t = -0.91$, $p = .363$), indicating a larger naming effect for the CP compared with the labial MP. The dorsal MP condition also had a significant effect on the intercept ($b = -0.33$, $t = -2.13$, $p = .003$) but not on the quadratic time term ($b = -2.14$, $z = -1.01$, $p = .310$), indicating a smaller and slower naming effect compared with the CP. Cohen's d indicated an effect size of $d = 0.552$ for the

---

[4] More stringent exclusion criteria would have led to the exclusion of 5 participants and 18% of their trials.

difference between CP and labial MP, and $d = 0.403$ for the difference between CP and dorsal MP.

In the language-specific time window, the percentage of looks to target was 56.7% in the CP condition, 48.6% in the labial MP condition, and 45.4% in the dorsal MP condition. In contrast to the Dutch sample, the alternative time window changes the results for the labial MP in the Japanese sample: The GCA neither showed a significant effect of the labial MP on the intercept ($b = -0.241$, $t = -1.44$, $p = .150$) nor on the quadratic time term ($b = -0.114$, $t = -0.54$, $p = .592$), indicating no difference in looking proportions between CP and labial MP. The dorsal MP condition had a significant effect on the intercept ($b = -0.372$, $t = -2.21$, $p = .027$) but not on the quadratic time term ($b = -0.018$, $t = -0.08$, $p = .932$), indicating a smaller naming effect compared with the CP. Cohen's d indicated an effect size of $d = 0.259$ for the difference between CP and labial MP, and $d = 0.491$ for the difference between CP and dorsal MP.[5]

The analysis on the difference score between prenaming and postnaming window showed an increase in target looks for the CP in both time windows: common, $m = 0.07$, $t(28) = 1.87$, $p = .072$; language specific, $m = 0.09$, $t(28) = 2.75$, $p = .010$. Consistent with the results of the GCA, for the labial MP a comparable naming effect was found in the language-specific, $m = 0.08$, $t(28) = 2.08$, $p = .047$, but not the common, $m = 0.04$, $t(28) = 1.29$, $p = .209$, time window. There was no naming effect for the dorsal MP: common, $m = -0.04$, $t(28) = -1.02$, $p = .318$; language specific, $m = 0.02$, $t(28) = 0.95$, $p = .353$.

Again, the outcomes of the NPST corroborate the findings of the previous analyses. Japanese toddlers looked less to the labial MP condition (40.7%) than to the CP condition (59.8%) in the time window between 984 and 1,168 ms after target word onset, although this effect failed to reach significance by a small margin, $t = 31.62$, $p = .055$. This weak MP effect is consistent with the fact that the GCA and the prepost analyses showed a difference between infants' reaction to the CP and labial MP for the general, but not the language-specific time-window, thus also providing only moderate evidence for toddlers' sensitivity to this MP. The effect for the dorsal MP reached significance, with significantly fewer looks to dorsal MPs (43.0%) than to CPs (58.0%) in the time window between 851 and 1,234 ms, $t = 62.20$, $p = .022$.

The parental vocabulary reports for Japanese are summarized in Table 4. The plosive-initial words Japanese toddlers understood most frequently were dorsal, followed by coronal and labial words. Words produced were also most frequently dorsal, followed by labial and coronal, consistent with previous studies (cf. de Boysson-Bardies & Vihman, 1991).

These results indicate commonalities between Dutch and Japanese toddlers in that they are both highly sensitive to the coronal-to-dorsal and less sensitive to the coronal-to-labial change. They also, however, reveal differences in that part of the analyses show statistically significant evidence for Japanese toddlers' sensitivity to the coronal-to-labial change, which was never the case for Dutch toddlers. Although the commonalities between language groups suggest that a common language-general bias underlies perceptual sensitivities, the differences in sensitivity to the coronal-to-labial change are more compatible with the predictions from the production account, which assumed that Dutch but not

Japanese toddlers would more readily accept labial-initial words as instances of coronal-initial words.

Taking the results on the dorsal MP into account, however, shows that the production account fails to predict the full pattern of results: Contrary to predictions, Japanese toddlers did not accept dorsal MPs as instances of coronal-initial words but showed a comparable response to Dutch toddlers: Toddlers from both language groups showed significant differences between CPs and dorsal MPs in both GCA and NPST. In summary, it therefore seems that none of the predictions presented in Table 1 can fully account for the results. As we will argue in the General Discussion, the outcomes are most compatible with an interaction between early, language-general biases and language experience.

General Discussion

The current study aimed to uncover the cause of toddlers' insensitivity to consonant mispronunciations in lexical tasks by cross-linguistically comparing toddlers' perception of the coronal-to-labial and coronal-to-dorsal changes. We found evidence for cross-linguistic commonalities such that toddlers from both language backgrounds showed high sensitivity to the coronal-to-dorsal change and reduced sensitivity to the coronal-to-labial change but also for cross-linguistic differences such that Japanese toddlers were more sensitive to the coronal-to-labial change than Dutch toddlers. These results suggest that perceptual insensitivities are specific to the coronal-to-labial change and that the observed insensitivities are in turn based on an interplay of early biases with language-specific input and production capacities. We will revisit this point after discussing the results in light of the predictions put forward in the introductory section.

Results in Light of Predictions

As discussed in the Results sections, neither early perceptual biases, underspecification, frequency, nor production can fully account for the present findings. If only an early coronal-labial perceptual bias (Tsuji et al., 2015) was causing Dutch toddlers' insensitivity to the coronal-to-labial change, this insensitivity should have been found with a similar strength in Japanese toddlers. Underspecification in turn would have predicted toddlers of neither language background to be sensitive to any of the mispronunciations.

The predictive power of the frequency account does not fare any better. Contrary to the prediction that change sensitivity would only be observed for the coronal-to-dorsal change in Japanese toddlers, sensitivity was found in all but the coronal-to-labial change in Dutch toddlers. In addition to the question of which type of frequency count is most relevant for toddlers it is an open question how large frequency differences would need to be to cause a perceptual asymmetry. Considering only those cases with relatively large frequency differences would predict a decreased sensitivity to the coronal-to-labial change only, since coronals are

---

[5] We again performed an additional analysis based on the more stringent exclusion criteria (see footnote 2). The results for the language-specific time-window remained qualitatively the same, whereas the common time-window analysis now also failed to show a difference between CP and labial MP ($b = -0.226$, $t = -1.46$. $p = .145$).

on average around twice as frequent as labials, 1.9 times in Dutch and 2.2 times in Japanese (while they are only 1.4 times as frequent as dorsals in Dutch, and dorsals are only 1.1 times as frequent as coronals in Japanese). Frequency might also only modulate the perception of speech sounds that are difficult to discriminate to begin with, and it is possible that the coronal-to-labial change (for which this difficulty is attested, cf. Tsuji et al., 2015), but not the coronal-to-dorsal change (for which no infant data are available) is difficult to discriminate initially. Indeed, frequency-dependent modulation of infants' native discrimination abilities has only been reported for speech sounds that were initially difficult to discriminate (e.g., Pons et al., 2012). However, even if we assume that frequency only affects the coronal-to-labial change, the frequency account fails: Because coronals are more frequent than labials in both languages, it would have predicted an equivalent lack of sensitivity in both Dutch and Japanese toddlers. In sum, it is therefore unlikely that simple frequency of exposure can account for the present results.

The production account is compatible with the diverging results on the coronal-to-labial change between Dutch and Japanese toddlers, but it would have predicted an insensitivity to the coronal-to-dorsal change in Japanese toddlers. Note, however, that this prediction was partly based on Japanese toddlers' early productive vocabulary, for which the difference between dorsal- and labial-initial words for Japanese is very small, thus potentially not being relevant for causing a perceptual insensitivity. In this light, the production account is most compatible with the current findings (but note that other sources we base our predictions on do provide evidence for a high occurrence of dorsals in Japanese toddlers' inventories, see Introduction).

In the following, we will discuss possible alternative explanations for the present results.

Perception of the Coronal-to-Dorsal Change

Because there is no published data on young infants' discrimination of the coronal-to-dorsal change, we turned to adult discrimination data to compare perceptual sensitivity to the coronal-to-labial and coronal-to-dorsal change. Data from both Dutch (Smits, Warner, McQueen, & Cutler, 2003) and Japanese (Saito, 1961) phoneme confusions suggest that discrimination of the coronal-to-labial change is more difficult than discrimination of the coronal-to-dorsal change.[6] Participants' confusion of a coronal with a labial plosive was consistently higher ([p] in place of [t]: 7.6% in Dutch, 7.0% in Japanese; [b] in place of [d]: 15.9% in Dutch, 2.0% in Japanese) than their confusion of a coronal with a dorsal plosive ([k] in place of [t]: 0.9% in Dutch, 1.3% in Japanese; [g] in place of [d]: 0.3% in Dutch, 1.5% in Japanese). Although infants' early discrimination abilities cannot straightforwardly be inferred from these adult data, they might at least indicate that the coronal-to-dorsal change is easier to perceive compared to the coronal-to-labial change for young infants to begin with. Speaking to this suggestion, Dutch and Japanese adults' perceptual patterns are comparable to those of young infants in Tsuji et al. (2015): They frequently confused coronals with labials and did so more than they confused labials with coronals ([t] in place of [p]: 1.3% in Dutch, 3.1% in Japanese; [d] in place of [b]: 1.5% in Dutch, 1.3% in Japanese). Therefore, it might be the case that the coronal-to-dorsal change is not as difficult to perceive for young infants as the

coronal-to-labial change, leading to the clear mispronunciation detection at 18 months of age. Future experiments on early discrimination in infants are, however, needed to back up this interpretation (and are currently underway in our lab).

Perception of the Coronal-to-Labial Change

A previous study has shown that both Dutch and Japanese prelexical infants are insensitive to the coronal-to-labial change (Tsuji et al., 2015). The current study showed that this early insensitivity is still observable in 18-month-old toddlers and a lexical task, indicating that early language-general biases keep influencing perceptual sensitivities. Nevertheless, we also observed a tendency for a language-dependent divergence in perceptual sensitivities later in development. The production account put forward in the introductory section predicted this divergence based on the fact that Dutch, but not Japanese, toddlers predominantly produce labial-initial words early on (de Boysson-Bardies & Vihman, 1991; Fikkert & Levelt, 2008), a pattern that is corroborated by norming data on early productions and also resembles infants' input types (see introductory section). The fact that labial-initial words are highly frequent in Dutch but not Japanese toddlers' early inventories could be the reason that Dutch but not Japanese toddlers treated labial MPs of previously learned coronal-initial words as acceptable variants. It is important that parental reports of toddlers' productive vocabularies in the current study reproduce this pattern: Although Dutch toddlers most frequently produced words with labial-initial plosives, Japanese toddlers produced words with dorsal-initial plosives (cf. Table 4).

Another language-specific factor that could possibly explain the divergence in the perception of the coronal-to-labial change in Dutch and Japanese toddlers is a difference in the distributional characteristics of coronals in Dutch versus Japanese. In Dutch, coronals undergo phonological processes like place assimilation which lead to surface variation. This context-dependent realization of syllable-final coronals as labials or dorsals could make Dutch coronals more variable compared to Japanese, and this variability could lead to fuzzier category boundaries in Dutch compared to Japanese toddlers. To validate this possibility, careful analyses of distributional characteristics of phonemes in both Dutch and Japanese combined with targeted experiments are necessary. More broadly, such detailed analysis would also provide insights into the kind of phonetic characteristics that do or do not influence infants' phoneme discrimination ability.

Conclusions and Outlook

The present data demonstrate that traces of an early cross-linguistic insensitivity to phonological detail persist across development but also shift toward more language-specific perceptual patterns. Further research needs to clarify at what point in development and on what level of representation Dutch and Japanese

---

[6] One difficulty in comparing data across these studies is that the Dutch study presented participants with gated fragments, whereas the Japanese study presented whole syllables varying signal-to-noise ratios. As it was difficult to determine which condition in the Japanese study was best comparable to the Dutch data, we report in the following the mean values across different conditions for the Japanese data.

toddlers' sensitivity to the coronal-to-labial change starts to diverge. In case distributional properties do impact on language-specific sensitivities, this divergence might be observed in phonetic discrimination within the first year of life. If toddlers' early word inventories are the critical difference, the divergence might only be observed once toddlers have a small receptive and productive vocabulary and possibly only in lexical tasks. For the latter possibility, it could even be the case that toddlers are able to perceive a difference in a discrimination task, but not in a lexical task. For fricative-plosive asymmetries, it has been demonstrated that 14-month-old infants were able to discriminate the plosive-to-fricative change but did not show sensitivity to the same change in a lexical context at 18 months of age (Altvater-Mackensen, & Fikkert, 2010; Altvater-Mackensen et al., 2014).

That toddlers are sensitive to the coronal-to-dorsal change is not surprising given that the majority of the literature demonstrates toddlers' sensitivity to phonological detail. They are, however, highly unexpected in the context of the underspecification literature. Not only would the account have predicted insensitivities to both types of tested changes but also have numerous adult studies documented insensitivities to both the coronal-to-labial and coronal-to-dorsal changes in German adult listeners (cf. Lahiri & Reetz, 2010, for an overview). It is an open question to what extent these findings can be reconciled with the results of the current study. One potential difference lies in the methods used to measure perceptual sensitivities. Adult listeners would certainly be able to detect a change in a simple and straightforward mispronunciation task, which is why the above-mentioned adult studies predominantly assessed responses in cross-modal priming or ERP studies. Thus, it might have been the case that sensitivity to the coronal-to-dorsal change was also reduced but to a lesser degree than sensitivity to the coronal-to-labial change and that the preferential looking task was not sensitive enough to capture it. This is, however, not a very likely explanation, as no adult study reports differences in strength of asymmetries between labial-coronal and coronal-dorsal contrasts. Further studies on the presence or absence of a language-general early asymmetry in the coronal-dorsal contrast, as well as tracking the development of perceptual asymmetries, might help to shed light on this question.

In sum, the current study demonstrated that toddlers are insensitive to phonological detail under certain circumstances: Early language-general perceptual insensitivities are maintained to a certain degree in lexical contexts but can be modulated by language-specific properties (such as distributional properties of phonemes in Japanese and the phonemes in Japanese toddlers' early vocabularies).

References

Altvater-Mackensen, N., & Fikkert, P. (2010). The acquisition of the stop-fricative contrast in perception and production. *Lingua, 120,* 1898–1909. http://dx.doi.org/10.1016/j.lingua.2010.02.010

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,* 149–178.

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

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. R package version 0.999999–0. Retrieved from http://cran.r-project.org/package=lme4

de Boysson-Bardies, B., & Vihman, M. (1991). Adaptation to language: Evidence from babbling and first words in four languages. *Language, 67,* 297–319. http://dx.doi.org/10.1353/lan.1991.0045

DeLoache, J. S., Chiong, C., Sherman, K., Islam, N., Vanderborght, M., Troseth, G. L., . . . O'Doherty, K. (2010). Do babies learn from baby media? *Psychological Science, 21,* 1570–1574. http://dx.doi.org/10.1177/0956797610384145

DePaolis, R. A., Vihman, M. M., & Keren-Portnoy, T. (2011). Do production patterns influence the processing of speech in prelinguistic infants? *Infant Behavior and Development, 34,* 590–601. http://dx.doi.org/10.1016/j.infbeh.2011.06.005

DePaolis, R. A., Vihman, M. M., & Nakai, S. (2013). The influence of babbling patterns on the processing of speech. *Infant Behavior and Development, 36,* 642–649. http://dx.doi.org/10.1016/j.infbeh.2013.06.007

Dijkstra, N., & 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 N. Danis, K. Mesh, & H. Sung (Eds.), *Proceedings of the 35th Annual Boston University Conference on Language Development* (Vol. 1, pp. 170–182). Somerville, MA: Cascadilla Press.

Edwards, J., & Beckman, M. E. (2008). Some cross-linguistic evidence for modulation of implicational universals by language-specific frequency effects in phonological development. *Language learning and development: the official journal of the Society for Language Development, 4,* 122–156. http://dx.doi.org/10.1080/15475440801922115

Fennell, C. T., & Waxman, S. R. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development, 81,* 1376–1383. http://dx.doi.org/10.1111/j.1467-8624.2010.01479.x

Fennel, C. T., van der Feest, S. V. H., & Spring, M. (2011). *Consonants asymmetries at 14 months: Implications of acoustic variation on early phonological acquisition.* Manuscript in preparation.

Fikkert, P. (2010). Developing representations and the emergence of phonology: Evidence from perception and production. In C. Fougeron, B. K. 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. (2008). How does place fall into place? The lexicon and emergent constraints in the developing phonological grammar. In P. Avery, E. Dresher, & K. Rice (Eds.), *Contrast in phonology: Perception and Acquisition* (pp. 231–270). Berlin, Germany: Mouton.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. *Nature Reviews Neuroscience, 5,* 831–843. http://dx.doi.org/10.1038/nrn1533

Lahiri, A., & Reetz, H. (2010). Distinctive features: Phonological underspecification in representation and processing. *Journal of Phonetics, 38,* 44–59. http://dx.doi.org/10.1016/j.wocn.2010.01.002

Li, F., Munson, B., Edwards, J., Yoneyama, K., & Hall, K. (2011). Language specificity in the perception of voiceless sibilant fricatives in Japanese and English: Implications for cross-language differences in speech-sound development. *The Journal of the Acoustical Society of America, 129,* 999–1011. http://dx.doi.org/10.1121/1.3518716

Mani, N., & Plunkett, K. (2010a). Twelve-month-olds know their cups from their keps and tups. *Infancy, 15,* 445–470. http://dx.doi.org/10.1111/j.1532-7078.2009.00027.x

Mani, N., & Plunkett, K. (2010b). In the infant's mind's ear: Evidence for implicit naming in 18-month-olds. *Psychological Science, 21,* 908–913. http://dx.doi.org/10.1177/0956797610373371

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods, 164,* 177–190. http://dx.doi.org/10.1016/j.jneumeth.2007.03.024

Mazuka, R., Igarashi, Y., & Nishikawa, K. (2006). *Input for learning Japanese: RIKEN Japanese Mother-Infant Conversation Corpus.* The Institute of Electronics, Information and Communication Engineers (Tech. Rep. No. 16), 11–15.

Mirman, D. (2014). *Growth curve analysis and visualization using R.* Boca Raton, FL: Chapman and Hall/CRC Press.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition, 98,* 13–30. http://dx.doi.org/10.1016/j.cognition.2004.10.005

Nazzi, T., & Bertoncini, J. (2009). Phonetic specificity in early lexical acquisition: New evidence from consonants in coda positions. *Language and Speech, 52,* 463–480. http://dx.doi.org/10.1177/0023830909336584

Ogura, T., & Watamaki, T. (2004). *Technical manual of the Japanese MacArthur Communicative Development Inventory: Words and Gesture.* Kyoto, Japan: Kyoto International Social Welfare Exchange Center.

Paradis, C., & Prunet, J. F. (Eds.). (1991). *The special status of coronals: Internal and external evidence (Phonetics and Phonology 2).* San Diego, CA: Academic Press.

Pons, F., Albareda-Castellot, B., & Sebastián-Gallés, N. (2012). The interplay between input and initial biases: Asymmetries in vowel perception during the first year of life. *Child Development, 83,* 965–976. http://dx.doi.org/10.1111/j.1467-8624.2012.01740.x

R Core Team. (2012). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.* Retrieved from http://www.R-project.org/

Ren, J., & Morgan, J. L. (2012). The devil in the details: Underspecification in infants' early lexical representations. *Proceedings of the 36th Annual Boston University Conference on Language Development* (pp. 500–511). Somerville, MA: Cascadilla Press.

Saito, S. (1961). Confusion matrices of Japanese speech sounds. *Electrical Communication Laboratory Technical Journal, 6.*

Schlichting, J., & Spelberg, H. L. (2002). *Lexilijst Nederlands, een instrument om de taalontwikkeling te onderzoeken bij Nederlandstalige kinderen van 15–27 maanden in het kader van vroegonderkenning: Manual* [Lexilijst Dutch, an instrument to assess language development of Dutch-learning children between 15–27 months for early identification: Manual]. Lisse, the Netherlands: Swets Test Publishers.

Shi, R., & Werker, J. F. (2001). Six-month-old infants' preference for lexical words. *Psychological Science, 12,* 70–75. http://dx.doi.org/10.1111/1467-9280.00312

Smits, R., Warner, N., McQueen, J. M., & Cutler, A. (2003). Unfolding of phonetic information over time: A database of Dutch diphone perception. *The Journal of the Acoustical Society of America, 113,* 563–574. http://dx.doi.org/10.1121/1.1525287

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388,* 381–382. http://dx.doi.org/10.1038/41102

Swingley, D. (2003). Phonetic detail in the developing lexicon. *Language and Speech, 46,* 265–294. http://dx.doi.org/10.1177/00238309030460021001

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental Science, 8,* 432–443. http://dx.doi.org/10.1111/j.1467-7687.2005.00432.x

Swingley, D. (2009). Onsets and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language, 60,* 252–269. http://dx.doi.org/10.1016/j.jml.2008.11.003

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition, 76,* 147–166. http://dx.doi.org/10.1016/S0010-0277(00)00081-0

Tsuji, S., Mazuka, R., Cristia, A., & Fikkert, P. (2015). Even at 4 months, a labial is a good enough coronal, but not vice versa. *Cognition, 134,* 252–256. http://dx.doi.org/10.1016/j.cognition.2014.10.009

van der Feest, S. V. H., & Fikkert, P. (in press). Learning to represent phonological contrasts: Evidence from perception and production. *Phonology.*

Van de Weijer, J. (1998). *Language input for word discovery.* MPI series in psycholinguistics 9.

Vihman, M., Kay, E., de Boysson-Bardies, B., Durand, C., & Sundberg, U. (1994). External sources of individual differences? A cross-linguistic analysis of the phonetics of mothers' speech to 1-yr-old children. *Developmental Psychology, 30,* 651–662. http://dx.doi.org/10.1037/0012-1649.30.5.651

Von Holzen, K., & Mani, N. (2012). Language nonselective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology, 113,* 569–586. http://dx.doi.org/10.1016/j.jecp.2012.08.001

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,* 1–30. http://dx.doi.org/10.1207/S15327078IN0301_1

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language, 59,* 114–132. http://dx.doi.org/10.1016/j.jml.2008.03.001

Yoshida, K. A., Fennell, C. T., Swingley, D., & Werker, J. F. (2009). Fourteen-month-old infants learn similar-sounding words. *Developmental Science, 12,* 412–418. http://dx.doi.org/10.1111/j.1467-7687.2008.00789.x

Zink, I., & Lejaegere, M. (2003). *N-CDI's: Korte vormen, Aanpassing en hernormering van de MacArthur Short Form Vocabulary Checklist van Fenson et al.* [N-CDIs: Short forms, adaption and renorming of the MacArthur Short Form Vocabulary Checklist of Fenson et al.]. Leuven, Belgium: Acco.