The Official Journal of the International Society on Infant Studies
Infancy

Infancy, ***18***(6), 1030–1052, 2013

ISSN: 1525-0008 print / 1532-7078 online
DOI: 10.1111/infa.12022

WILEY Blackwell

The Impact of Mispronunciations on Toddler Word Recognition: Evidence for Cascaded Activation of Semantically Related Words from Mispronunciations of Familiar Words

Nicole Altvater-Mackensen
Research Group Early Social Development
Max Planck Institute for Human Cognitive and Brain Sciences
Research Group Language Acquisition
Georg-August-University Göttingen

Nivedita Mani
Research Group Language Acquisition
Georg-August-University Göttingen

While the specificity of infants' early lexical representations has been studied extensively, researchers have only recently begun to investigate how words are organized in the developing lexicon and what mental representations are activated during processing of a word. Integrating these two lines of research, the current study asks how specific the phonological match between a perceived word and its stored form has to be in order to lead to (cascaded) lexical activation of related words during infant lexical processing. We presented German 24-month-olds with a cross-modal semantic priming task where the prime word was either correctly or incorrectly pronounced. Results

Correspondence should be sent to Nicole Altvater-Mackensen, Forschungsgruppe Frühe soziale Entwicklung, Max-Planck-Institut für Kognitions- und Neurowissenschaften, Stephanstrasse 1a, D-04103 Leipzig, Germany. E-mail: altvater@cbs.mpg.de

indicate that correct pronunciations and mispronunciations both elicit similar semantic priming effects, suggesting that the infant word recognition system is flexible enough to handle deviations from the correct form. This might be an important prerequisite to children's ability to cope with imperfect input and to recognize words under more challenging circumstances.

Around their second birthday, toddlers already know a few hundred words (e.g., Fenson et al., 1994). There is ample evidence suggesting that children have detailed knowledge about these words. For instance, mispronunciation detection studies show that toddlers are sensitive to subtle one-feature mispronunciations of consonants and vowels in well-known as well as recently learned words (Bailey & Plunkett, 2002; Mani & Plunkett, 2007; Swingley, 2009; Swingley & Aslin, 2000; among others); and they appear to be sensitive to the size of a mispronunciation, showing greater sensitivity to larger mispronunciations relative to smaller mispronunciations (Mani & Plunkett, 2010a; White & Morgan, 2008). These findings have been taken as evidence that toddlers possess highly specified phonological representations of words. Still, there are a number of studies suggesting that not all mispronunciations are detected equally well. For instance, mispronunciations in voicing or vowel height might be hard for toddlers to detect (Mani, Coleman, & Plunkett, 2008; Mani & Plunkett, 2010b; Van der Feest, 2007). Some studies also find asymmetries in toddlers' sensitivity to place or manner of articulation differences (Altvater-Mackensen, 2010; Van der Feest, 2007) as well as differences in toddlers' sensitivity to vowel and consonant mispronunciations (Havy & Nazzi, 2009; Nazzi, 2005; Nazzi, Floccia, Moquet, & Butler, 2009).

Alongside these differences in sensitivity to certain sound contrasts, speaker-specific characteristics have also been shown to influence word recognition in toddlers. While toddlers have no difficulties in recognizing familiar words spoken in their native dialect, they have problems to do so when words are uttered with a non-native accent or in an unfamiliar dialect (Best, Tyler, Gooding, Orlando, & Quann, 2009; Schmale, Hollich, & Seidl, 2011). Yet, brief exposure to a particular dialect (e.g., hearing a story read out in this dialect) can help toddlers adapt to the dialect-specific characteristics and to subsequently recognize words spoken in this dialect (Schmale, Christia, & Seidl, 2012; Van Heugten & Johnson, 2013). Similarly, White and Aslin (2011) report that toddlers tolerate 'mispronunciations' that follow a vowel shift that they have heard in a naming context before (i.e., mispronunciations in keeping with the previously familiarized dialect of the speaker), but do not tolerate mispronunciations not in keeping with the familiarized dialect. Taken together, these studies suggest that toddlers can adapt to systematic variation in the input, such as dialectal variation, and that this enables them to recognize words that follow a speaker-specific rather than a canonical pronunciation (e.g., Schmale et al., 2012; White & Aslin, 2011).

While the above-mentioned studies help to broaden our knowledge about the specificity of children's early lexical representations and their ability to cope with systematic variability, such as dialectal variation, they do not necessarily inform us of the extent to which lexical processing is disrupted by a deviant pronunciation. Although accent studies suggest that 'mispronunciations' do not prevent word recognition when they follow a predictable rule, they do not inform us about the impact of unpredicted, unsystematic variation, such as spontaneous speech errors. Does a mispronunciation prevent lexical activation of the target label? Or does the phonological overlap between target label and mispronunciation still lead to lexical activation of the target label? In other words, do children still recognize a word when it is mispronounced?

Some mispronunciation detection studies report that children fixate the target above chance not only after hearing its correctly pronounced label but also after hearing a mispronunciation (Swingley & Aslin, 2000, 2002; but see Ballem & Plunkett, 2005; Mani & Plunkett, 2007; White & Morgan, 2008). Looking at the time course of target fixations, it seems that children might eventually fixate the target after hearing a mispronounced or accented version of its label (e.g., Creel, 2012; Swingley, 2009). This might be taken to suggest that they activate the correct or standard pronunciation of the target label – albeit later – even upon hearing a deviant pronunciation of this label (henceforth mispronunciation) and that they treat it as a poor realization of the target label (rather than as a novel, unfamiliar word). Yet, it would also be reasonable to suggest that children searching for the mispronunciation's referent find that the perceived word does not refer to the distracter and that fixations to the target are reinforced by the sublexical or phonological overlap between mispronunciation and target label. This might especially be so, given that pre-exposure to a familiar target image before naming allows the child to internally generate the target label (e.g., Mani & Plunkett, 2010c; Meyer, Belke, Telling, & Humphreys, 2007). The child may then identify the target picture as the referent for the heard (mispronounced) label by matching this sound with the formerly retrieved label, rather than the mispronunciation leading to lexical activation of the target word. In other words, the image leads to the activation of the target word and the child recognizes that the phonological properties of the heard mispronounced label match the stored information about the target label more than the stored information about the distracter label. Note that toddlers tend to look at the distracter image more reliably upon hearing a mispronunciation when the distracter is unfamiliar (i.e., when the child has no label for the distracter) compared with when the distracter is familiar. This is in keeping with our suggestion that isolating the target referent involves matching the phonological properties of the heard label with the phonological properties of the labels associated with the depicted referents (see also discussion in White & Morgan, 2008). Thus, mispronunciation detection studies can help to reveal how specific phonological representations are, but they do not necessarily inform us of the extent to which a mispronunciation can lead to lexical activation of the correct pronunciation of the target label.

Yet, making the connection between a mispronunciation and the correct form of the target label might be a crucial prerequisite to cope with imperfect input. It is only when the word recognition system is flexible enough to handle deviations from the correct form that the child will be able to recognize words under more challenging circumstances, such as when noise masks parts of the signal, or when the speaker's articulation is sloppy or prone to errors. In addition, the visual environment in natural situations might be more crowded relative to controlled experimental settings where the child is faced with only two possible referents (but see Yu and Smith (2012) for a different view on the ambiguity problem). Under such circumstances, it seems even more important to activate the correct target label upon hearing its phonological form (which might not always be perfectly accurate), as preactivation of the labels of all visible objects might be less likely to help in disambiguating the correct referent.

One way to examine whether mispronunciations lead to activation of the correct pronunciations of a word- avoiding potential reinforcement of the target word from a picture- is to test whether mispronunciations can lead to activation of words lexically related to the target. Researchers have only recently begun to investigate how words are organized in the developing lexicon and what mental representations are activated during processing of a word. All these studies used known words, that is, correctly pronounced labels or images of words that were familiar to the child. They reveal how processing of a familiar word involves the activation of lexically related words, but do not necessarily inform us about the mechanisms underlying lexical activation triggered by unknown words or mispronunciations.

Priming studies indicate that, upon hearing a familiar word, toddlers activate other words semantically (Arias-Trejo & Plunkett, 2009) as well as phonologically related (Mani & Plunkett, 2010c, 2011) to the heard word. Arias-Trejo and Plunkett (2009) showed that 2-year-olds' recognition of a target word like *cat* is facilitated by previous exposure to a semantically related prime, such as *dog*. This suggests that words that have similar meaning are interconnected in the child's lexicon and that activation of a word such as *cat* involves activation of words semantically related to *cat*, such as *dog*. Indeed, recent studies have also shown that phonological overlap between prime and target labels similarly influences target recognition in 2-year-olds (Altvater-Mackensen & Mani, under review; Mani & Plunkett, 2010c, 2011).[1] Furthermore, activation seems to cascade from phonological to semantic levels of representation with phonologically related words triggering activation of other semantically related words (Huang & Snedeker, 2010; Mani, Durrant, & Floccia, 2012). For instance, Mani et al. (2012) found that 2-year-olds' recognition of a target word like *shoe* is facilitated by previous exposure to the prime *clock*. As *clock* and *shoe* are neither semantically nor phonologically related, this suggests a cascaded pattern of activation: presentation of the prime *clock* leads to activation of the similar-sounding subprime *sock*, which in turn leads to activation of its semantic associate *shoe*. These findings strongly resemble findings of cascaded activation in adults (Huettig & McQueen, 2007; Yee & Sedivy, 2006) and support models of word recognition assuming that processing of a word involves activation of a set of possible word candidates (Luce & Pisoni, 1998; Marslen-Wilson, 1987; McClelland & Elman, 1986; Norris, 1994).

What remains unclear, and what will be addressed in the current study, is whether (cascaded) lexical activation in infants requires an accurate match between a perceived word and its stored form, that is, a correct pronunciation of a familiar word. Put differently, we ask whether mispronunciations can also lead to (cascaded) lexical activation. As stated above, previous studies do not answer this question because they tested lexical activation triggered by familiar word primes and not nonwords. Specifically, the current experiment investigates whether hearing a mispronunciation of a word (e.g., the mispronunciation *gat* for *cat*) leads to cascaded activation of words semantically related to the correct pronunciation (e.g., *dog*). Research with adults suggest that deviant pronunciations can indeed prime words semantically related to the correct pronunciation of a word, such as the nonword *gat* priming *dog* through the correct pronunciation *cat* (Milberg, Blumstein, & Dworetzky, 1988; Misiurski, Blumstein, Rissman, & Berman, 2005). If we find a similar pattern of cascaded activation upon hearing a mispronunciation in toddlers, this would strongly suggest that mispronunciations do not completely disrupt word recognition. Rather, it

[1]The absence of (semantic) priming in toddlers younger than two years of age does not mean that lexical representations have no organization at the phonological level. Indeed, children's early vocabularies tend to have dense phonological neighborhoods, suggesting an influence of phonological similarity on word learning (e.g., Stokes, 2010). Furthermore, Mani and Plunkett (2011) observed that phonological overlap facilitates word recognition in toddlers as young as 18 months of age (but see their discussion for the suggestion that these effects are sublexical).

would indicate that children activate the stored form of a word even when being presented with deviant pronunciations of the same word.

Note that using a (phono-) semantic priming task has a major advantage over using a phonological priming task. For example, improved recognition of the word *cat* given previous exposure to the mispronunciation *gat* (similar to mispronunciation studies to date) might solely rely on prelexical activation of the phonemes shared between the mispronunciation and the correct pronunciation (see also Radeau, Morais, & Segui, 1995; Slowiaczek & Hamburger, 1992). In other words, *cat* might be primed by its mispronunciation *gat* because hearing *gat* preactivates the phonemes/a/and/t/. This phonological-level activation can then sustain recognition of *cat*. This scenario does not, however, necessarily involve lexical-level activation of the word *cat* upon hearing *gat*. Activation of the semantically related word *dog* from the mispronunciation *gat*, however, provides evidence for the lexical retrieval of *cat* (and other words semantically related to cat) upon hearing *gat*. This is because the mispronunciation *gat* itself neither overlaps phonologically nor semantically with *dog*. Only the subprime *cat* (i.e., the correct form) can mediate activation of *dog*. Our design therefore closely matches the reasoning in Mani et al. (2012), while using mispronunciations instead of phonologically related, familiar words. If the mispronunciation leads to activation of the correct form – similar to a familiar word leading to activation of phonologically related words – we expect to find similar results as Mani et al. (2012). If mispronunciations do not lead to activation of the correct form – and, being nonwords, to activation of no other lexical form either – we do not expect (phono-) semantic priming to occur.

The question in how far mispronunciations lexically activate the correct form of a word is not trivial for at least two reasons. First, as pointed out above, previous mispronunciation detection studies do not necessarily inform us about the lexical activation of the target word upon hearing a mispronunciation. Evidence for such activation would, however, speak to the robustness of the infant word recognition system (i.e., how well it copes with imperfect input). Although infant-directed speech is usually characterized by careful articulation (e.g., Snow & Ferguson, 1977), the child will eventually have to be able to recognize words under more challenging circumstances (e.g., amidst background noise or sloppily articulated). This differs from their ability to cope with dialectal variation, as accents alter words in a predictable way, whereas mispronunciations are unpredictable. Second, examining the mental representations activated by mispronunciations of words can help distinguish between models of word recognition that put special emphasis on the beginnings of words, such as Cohort (Marslen-Wilson, 1987), and continuous models of word recognition, such as TRACE (McClelland & Elman, 1986) or NAM (Luce & Pisoni, 1998), because both types of models make different predictions regarding the impact of initial mispronunciations on lexical activation. Although initial mispronunciations might hinder word recognition in any of these models, Cohort (Marslen-Wilson, 1987) is the only model predicting that initial mispronunciations will completely block lexical access. Thus, if word beginnings are indeed of particular relevance for word recognition, then an initial mispronunciation like *gat* should block activation of (the correct form) *cat* and therefore not prime *dog*. If, however, overall phonological overlap is crucial for lexical activation, *gat* might still lead to activation of *cat* and thereby prime *dog* (see also Marslen-Wilson & Zwitserlood, 1989).

METHOD

Participants

Forty-four German 24-month-olds (19 boys) participated in the experiment. Their ages ranged from 23; 07 (months; days) to 25; 09, mean age 24; 07. Nine additional children (6 boys) were tested but excluded because of fussiness (2), parental interference (2) or because the child was talking throughout the whole experiment (5). Participants came from a sample of families who responded to an invitation letter sent to all families with infants living in the area. Participants were rewarded with a book or a t-shirt.

Stimuli

Table 1 details the set of stimuli presented to the children. Target images were presented with yoked distracter images whose label shared grammatical gender with the target label but were neither phonologically nor semantically related to the target (e.g., target *Kuh* ‘cow’ – distracter *Gabel* ‘fork’). Targets were preceded by auditory primes that were unrelated, semantically related or phono-semantically related to the target. Primes were neither semantically nor phonologically related to the distracter labels. Unrelated primes were phonologically and semantically unrelated to the target (e.g., prime *Buch* ‘book’ – target *Kuh* ‘cow’). Semantic primes, henceforth correctly pronounced primes, came from the same semantic category as the target (10 of 12 pairs, e.g., prime *Schaf* ‘sheep’ – target *Kuh* ‘cow’) or were associatively related to the target (2 of 12 pairs, e.g., prime *Baum* ‘tree’ – target *Vogel* ‘bird’). Phono-semantic primes, henceforth mispronounced primes, were mispronunciations of the same words used as correctly pronounced primes (e.g., prime *Taf* (mispronunciation of *Schaf* ‘sheep’) – target *Kuh* ‘cow’). Mispronunciations involved the change of the manner or the place of articulation of the prime’s initial consonant, a mispronunciation

TABLE 1
Stimulus Set

	Prime			
Unrelated	*Semantic (correct)*	*Phono-Semantic (mispronounced)*	*Target*	*Distracter*
Stein 'stone'	Ball 'ball'	Gall/Wall	Puppe 'doll'	Sonne 'sun'
Schrank 'closet'	Baum 'tree'	Gaum/Waum	Vogel 'bird'	Löffel 'spoon'
Auge 'eye'	Bett 'bed'	Gett/Wett	Decke 'blanket'	Tomate 'tomato'
Rutsche 'slide'	Birne 'pear'	Girne/Wirne	Apfel 'apple'	Pullover 'jumper'
Mund 'mouth'	Bus 'bus'	Gus/Wus	Auto 'car'	Telefon 'phone'
Fahrrad 'bike'	Bein 'leg'	Gein/Wein	Arm 'arm'	Hase 'rabbit'
Traktor 'tractor'	Finger 'finger'	Schinger/Pinger	Zeh 'toe'	Kuchen 'cake'
Mond 'moon'	Fisch 'fish'	Schisch/Pisch	Ente 'duck'	Hose 'trousers'
Eimer 'bucket'	Fuss 'foot'	Schuss/Puss	Hand 'hand'	Katze 'cat'
Elefant 'elephant'	Schuh 'shoe'	Fuh/Tuh	Jacke 'jacket'	Uhr 'clock'
Buch 'book'	Schaf 'sheep'	Faf/Taf	Kuh 'cow'	Gabel 'fork'
Affe 'ape'	Schal 'scarf'	Fal/Tal	Mütze 'hat'	Windel 'nappy'

that is easily detected by 24-month-olds (e.g., Swingley & Aslin, 2000, 2002). The mispronunciations changed the semantic primes into nonwords in all except three cases where mispronunciations were words that 24-month-olds are unlikely to know (*Wall* 'bank', *Wein* 'wine', *Tal* 'valley').

Visual stimuli (26.5 cm $\times$ 36 cm) consisted of photographs of the target and distracter objects against a light gray background. Audio stimuli consisted of the prime and target labels, spoken by a female native speaker of German, using child directed speech. Audio stimuli were digitally recorded in a quiet room with a sampling rate of 44,100 Hz and volume matched after recording using audio editing software. Mean duration, pitch and pitch range of primes did not significantly differ across conditions (*ps* > .1) (correctly pronounced primes: mean duration = 623 ms, mean f_0 = 263 Hz, mean f_0 range = 176 Hz; mispronounced primes: mean duration = 617 ms, mean f_0 = 263 Hz, mean f_0 range = 169 Hz; unrelated primes: mean duration = 671 ms, mean f_0 = 258 Hz, mean f_0 range = 189 Hz). The target label was reused across conditions (mean duration = 623 ms, mean f_0 = 260 Hz, mean f_0 range = 189 Hz).[2]

[2]Unrelated primes tended to be slightly longer than correctly or mispronounced semantic primes because there were more bisyllabic unrelated primes included than bisyllabic related primes. However, the unrelated condition served as control condition for both correctly and mispronounced primes. As correctly and mispronounced primes did not differ in length or pitch and as target labels were the same across conditions, no differences in the size or timing of the priming effect between correctly and mispronounced primes should arise from the acoustic characteristics of the stimuli.

Procedure

Prior to their visit, parents filled out a subset of the *Fragebogen zur frühkindlichen Entwicklung FRAKIS* (Szagun, Stumper, & Schramm, 2009), a standardized German communicative development inventory, to ensure that children knew the items used in the experiment. After a short play session during which parents were informed about the procedure of the experiment, children were seated on their parent's lap in a dimly lit, quiet experimental room, facing a TV screen (92 cm $\times$ 50 cm) at a distance of 100 cm from the screen. Two cameras mounted directly above where the images would appear on the TV screen recorded children's eye movements during the experiment. Synchronized signals from the cameras were routed via a digital splitter to record two separate time-locked images of the child. Auditory stimuli were presented via loudspeakers that were located above the screen. Stimuli were presented using the Look software (Meints & Woodford, 2008). Parents wore headphones playing music intermixed with speech during the experiment and were instructed to interact as little as possible with their child and to avoid pointing to the screen or naming the objects.

Each child was presented with three blocks of twelve trials. The stimulus set was counterbalanced across blocks so that each child saw every target and distracter pair once per block and so that each block contained four trials with an unrelated prime, four trials with a correctly pronounced prime, and four trials with a mispronounced prime. Prime-target pairs were distributed across blocks such that each target appeared with a different prime in each block. The order of blocks was counterbalanced across children. In between blocks, the children participated in two other preferential looking experiments that were unrelated to the current task to avoid effects of target repetition. One experiment was a naming study in which children saw pictures of familiar objects next to each other with one being (correctly) named. The other experiment was a word-learning study in which the child was first presented with a novel object and label; learning was then tested by presenting pictures of two objects next to each other and labeling one of them. Again, labeling included no mispronunciations. It was counterbalanced across children which study was presented after the first and second block of the current experiment. Both studies used images and labels that were not part of the stimulus set of the current experiment (i.e., none of the presented images or labels was repeated across experiments). Blocks of the current experiment were separated by approximately 3.5 minutes by the intervening studies.

During each trial, the child first saw a white cross in the middle of the TV screen against a black background for 1,500 ms. The auditory stimulus was

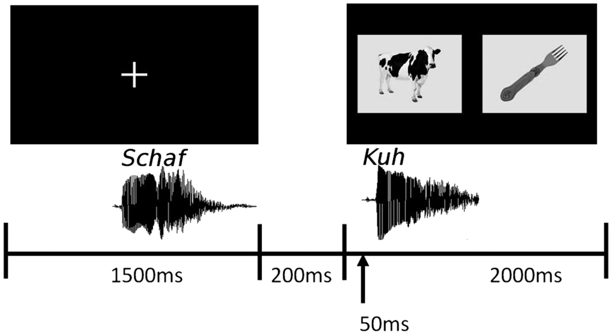

Figure 1 Schematic of the trial structure with stimulus examples. Note that original pictures were coloured. In the correct pronunciation condition, for the spoken target word *Kuh* 'cow', the picture of a cross was accompanied by the label *Schaf* 'sheep' (prime), followed by the pictures of a *Kuh* 'cow' (target) and a *Gabel* 'fork' (distracter) next to each other and the presentation of the target label *Kuh* 'cow'. In the mispronunciation condition, children hear *Taf*, in the place of *Schaf* 'sheep'.

presented such that the prime word's offset was aligned to the disappearance of the cross at 1,500 ms. This was followed by a silent, 200 ms interstimulus interval where the screen remained black. About 1,700 ms into the trial, target, and distracter images appeared side by side on the screen separated by 5.5 cm, against a black background. This configuration of stimuli has proved successful in tapping into phonological and semantic priming effects in word recognition by 24-month-old toddlers (Arias-Trejo & Plunkett, 2009; Mani & Plunkett, 2011). Fifty ms following the appearance of the target and distracter images on-screen (i.e., 1,750 ms into the trial) children were presented with the auditory target label. Target and distracter images stayed on-screen until the trial ended (i.e., 2,000 ms after the onset of the target label). In sum, each trial lasted 3,750 ms. Figure 1 presents a schematic of the trial structure with an example for a stimulus display. The side on which target and distracter appeared as well as the order of trial presentation was randomized within blocks.

Data analysis

Children's looking behavior was analyzed using a digital video scoring system. A trained coder indicated for each 40 ms frame of the video whether the child was looking to the left, to the right, in the middle, or away from the screen. The coder was blind to target location and trial type. A second trained coder recoded 10% of the videos to check the reliability of the coding. The codes provided by the two coders were reliably correlated (r = .98). The coding output was aligned with information about side of target, target word onset, and trial type to determine the amount of time that children spent looking at the target (T) and at the distracter (D) throughout each trial. The proportion of target looking, PTL = T/(T + D), was calculated across the time window from 360 ms after target word onset until the end of the trial. This is standard in the infant literature, because eye movements before 360 ms are not likely to be made in response to the auditory perception of the target word (e.g., Canfield et al., 1997). Trials in which children did not know the prime and/or the target label (according to individual communicative inventory reports filled out by the parents of the children) were excluded from analysis (148 trials = 11% of data). Furthermore, only those trials in which children looked at least once at target and/or distracter postnaming were included in the final analysis (21 trials = 2% of data excluded). This criterion controlled that only those trials where children were attending to the stimuli on screen are included in the analysis. The minimum looking time in these trials was 120 ms, and infants fixated the screen less than one second in a marginal number of trials (correctly pronounced prime: 18 trials; mispronounced prime: 14 trials; unrelated prime: 16 trials). Mean looking time at the screen during the window of analysis was 1.5 sec (i.e., infants fixated the screen for 91.5% of the time). Attention was similar across conditions (mean looking time correctly pronounced prime: 1,480 ms, SE 12.7; mispronounced prime: 1,486 ms, SE 11.1; unrelated prime: 1,482 ms, SE 12.2).

All other trials were included in the analysis, as long as the child provided data for least 50% of trials (6 children excluded). The final data set included data from 38 children, and a total of 1,051 trials that were evenly distributed across conditions (correctly pronounced prime: 348 trials; mispronounced prime: 352 trials; unrelated prime: 351 trials). The data were aggregated by condition (unrelated, correctly, or mispronounced prime) and block (one, two, three) for each child in the subject-wise analysis and for each target item in the item-wise analysis.

RESULTS

Figure 2 displays the mean total looking time to target and distracter as well as the time that children did not fixate any of the two pictures for unrelated, correctly pronounced (CP) and mispronounced (MP) semantic prime trials in the critical time window from 2,110 ms onwards until the end of the trial at 3,750 ms (i.e., 360 to 2,000 ms after target word onset). As can be seen from the graph, overall looking time was similar across conditions, but the ratio between target and distracter looking appears to be modulated by the prime. Figure 3 shows the mean proportion of target

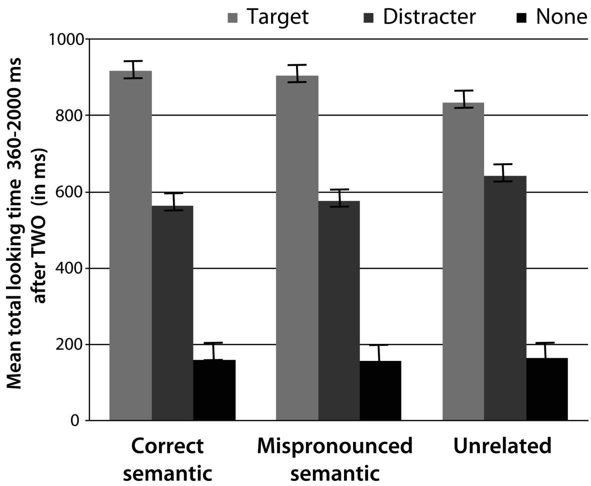

Figure 2 Mean total looking time (ms) to target, distracter or none of the images in correctly pronounced, mispronounced and unrelated prime trials in the window of analysis 360 to 2,000 ms after target word onset (TWO) (error bars indicate ± 1 SE).

looking in the same time window for all conditions (the pattern of results did not change when using total looking time to target rather than proportion of target looking as dependent measure). As can be seen from the graph, the proportion of target fixations is higher for CP- and MP-trials than for unrelated trials, while proportion of target looking in CP- and MP-trials seems to differ only marginally.

To examine whether children's looking behavior varied across the different blocks, we conducted a repeated measures ANOVA with condition (correctly pronounced, mispronounced, or unrelated prime) and block (one, two, three) as a within-subjects factor. The analysis showed a main effect of condition (subject analysis: $F(2, 36) = 5.486$, $p = .005$, $\eta_p^2 = .05$; item analysis: $F(2, 10) = 3.933$, $p = .024$, $\eta_p^2 = .11$), but no main effect of block ($p > .1$) and no interaction between condition and block ($p > .9$). This suggests that children's looking behavior was influenced by the nature of the prime, but did not differ across blocks. We therefore collapsed the data across blocks for further analysis.

A repeated measures ANOVA with condition (correctly pronounced, mispronounced, or unrelated prime) as a within-subjects factor showed a main effect of condition (subject analysis: $F(2, 36) = 8.769$, $p = .001$, $\eta_p^2 = .33$; item analysis: $F(2, 10) = 5.647$, $p = .023$, $\eta_p^2 = .53$). Planned post hoc two-tailed t-tests revealed that children look longer at the target after hearing a correctly pronounced semantically related prime relative to an

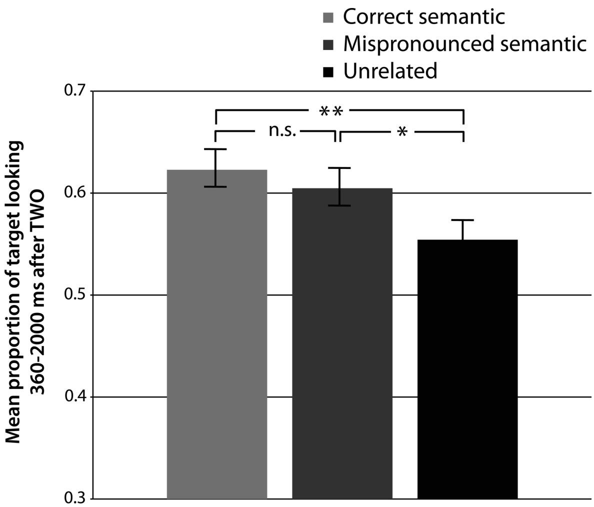

Figure 3 Mean proportion of target looking in correctly pronounced, mispronounced and unrelated prime trials in the window of analysis 360 to 2,000 ms after target word onset (TWO) (error bars indicate ± 1 SE).

unrelated prime (subject analysis: $t(37) = 4.240$, $p < .001$, $d = .82$; item analysis: $t(11) = 3.512$, $p = .005$, $d = .89$) and that they look longer at the target after hearing a mispronounced semantically related prime relative to an unrelated prime (subject analysis: $t(37) = 2.428$, $p = .020$, $d = .53$; item analysis: $t(11) = 2.056$, $p = .064$, $d = .49$). Thus, both correctly (e.g., *Schaf*, 'sheep') and mispronounced primes (e.g., *Taf*, mispronunciation of *Schaf*) facilitated recognition of a target (e.g., *Kuh*, 'cow') semantically related to the correct pronunciations of the prime.[3] There was, however, no difference in target looking following a correctly pronounced semantically related prime compared with a mispronounced semantically related prime (*ps* > .1). This suggests that target word recognition was equally facilitated by correctly and mispronounced primes.

To explore whether children recognized the target in the different priming conditions, we compared the proportion of target fixations against

[3]Arias-Trejo and Plunkett (2009) discuss that associatively related primes (such as *tree* and *bird*, which often cooccur but do not share semantic features) lead to larger priming effects than semantically related primes in infants. To ensure that our results are not driven by the two associatively related primes, we reran all analyses excluding these two primes. The only change in the pattern of results is the difference in looking time after hearing a mispronounced prime relative to after hearing an unrelated prime in the item analysis ($t(9) = 1.681$, $p = .13$). There were no other differences in the pattern of results.

chance (=.5). One-sample *t*-tests showed that target fixations were above chance for targets following a correctly pronounced prime (subject analysis: $t(37) = 8.663$, $p < .001$, $d = 1.41$; item analysis: $t(11) = 6.587$, $p < .001$, $d = 1.90$), a mispronounced prime (subject analysis: $t(37) = 7.340$, $p < .001$, $d = 1.19$; item analysis: $t(11) = 5.202$, $p < .001$, $d = 1.50$), and an unrelated prime (subject analysis: $t(37) = 3.401$, $p = .002$, $d = .55$; item analysis: $t(11) = 2.893$, $p = .015$, $d = .83$), suggesting that toddlers identified the named target, irrespective of whether this followed a related or an unrelated prime. We suggest that the labeling of the target leads to target fixations in all conditions, that is, irrespective of the prime. In related conditions, target fixations may be enhanced by previous activation of the target through the prime. In unrelated conditions, the prime does not provide such a benefit, but that does not imply that unrelated primes must harm target recognition.

Visual inspection of the proportion of toddlers' target fixations over the time course of the trial, however, suggests that looking behavior is modulated by the accuracy of the prime's pronunciation. Figure 4 shows the proportion of children's fixations at the target image every 40 ms in the trial separately for unrelated, correctly pronounced and mispronounced prime trials. As can be seen from the graph, children fixated the target more after hearing a correctly pronounced semantically related prime relative to an unrelated prime, while target fixations following a mispronounced semantically related prime lie in between.

To further analyze possible differences in the time course between correctly pronounced (CP) and mispronounced (MP) trials, we used the nonparametrical statistical test to identify time periods in which CP- and MP-trials differ from unrelated trials (Maris & Oostenveld, 2007; for application of the nonparametrical test in a preferential looking paradigm see Von Holzen & Mani, 2012). The nonparametrical statistical test compares the observed *t*-statistics testing differences between the time courses against a (large) sample of random permutations of the data. For each permutation, the data points are randomly assigned to one condition, and the *t*-statistic of the difference between conditions is computed. After 1,000 shuffles, the *t*-statistics of the actual data set are compared with the *t*-statistics of the random permutations to see whether the observed result falls within the expected range of *t*-statistics (created by the random permutations) or not. The resulting Monte Carlo *p*-value provides a measure of how likely it is that observed time course differences are due to chance (similar to standard *p*-values). Comparisons using this time course analysis revealed that CP-trials and unrelated trials deviated from each other between 3,160 and 3,720 ms (i.e., 1,410 to 1,970 ms after target word onset: cluster *t*-statistic = 30.67, Monte Carlo

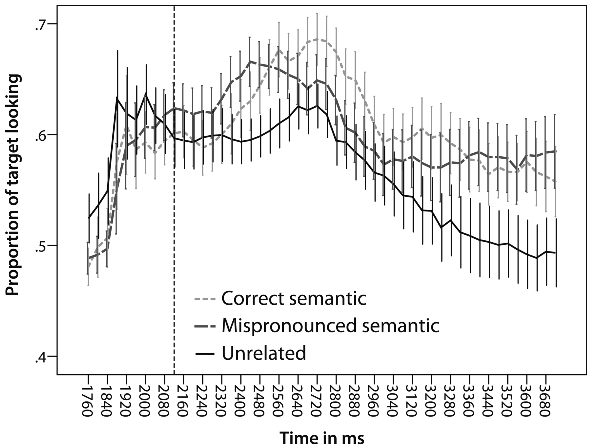

Figure 4 Time course graph showing the mean proportion of target fixations in correctly pronounced, mispronounced and unrelated prime trials for every 40 ms interval after the appearance of target and distracter picture (i.e., from 1,750 ms into the trial). The vertical striped line marks the onset of the window of analysis 360 ms after target word onset (i.e., 2,110 ms into the trial). Error bars indicate ± 1 SE.

p = .09) with enhanced target looking in CP-trials and that MP-trials and unrelated trials deviated from each other between 3,360 and 3,720 ms (i.e., 1,610 to 1,970 ms after target word onset: cluster t-statistic = 28.66, Monte Carlo p = .08) with enhanced target looking in MP-trials. The time course analysis indicates that the naming effect in MP-trials was 200 ms delayed compared with the naming effect in CP-trials (i.e., MP-trials started to differ from unrelated trials 200 ms later than CP-trials), suggesting that target recognition was indeed hindered by the mispronunciation.

DISCUSSION

The current study investigated the extent to which hearing a mispronunciation of a familiar word disrupts recognition of the word and influences subsequent cascaded activation of words semantically related to the correct pronunciation. Our results suggest that mispronunciations lead to cascaded activation of words semantically related to the correct pronunciations. Using the example from earlier, *gat* (a mispronunciation of *cat*) primes recognition of *dog*. This strongly suggests that the mispronunciation *gat* activates the phonologically related correct pronunciation *cat*, which in turn primes recognition of the semantically related word *dog*. The observed pattern is similar to findings of cascaded activation upon hearing correctly pronounced words in children (Huang & Snedeker, 2010; Mani et al., 2012) and adults (Huettig & McQueen, 2007; Milberg et al., 1988; Misiurski et al., 2005; Yee & Sedivy, 2006). Our results extend the previous studies with toddlers by showing that not only correctly pronounced words, but also mispronunciations can lead to cascaded lexical activation.

Similar effects of cascaded activation through mispronunciations have been reported for adults, for example by Marslen-Wilson and Zwitserlood (1989) who found that Dutch listeners show facilitated recognition of *bij* ‘bee’ when primed with *woning* (a mispronunciation of *honing* ‘honey’). Note that there are also several studies that show phonological priming effects for mispronunciations in adults, like *rater* priming recognition of the correct pronunciation *water* (e.g., Connine, Blasko, & Titone, 1993; Lukatela, Eaton, Lee, & Turvey, 2001). These studies report facilitated target word recognition following presentation of a mispronunciation of the target word. However, these effects might result from the phonological overlap between the mispronounced prime and the correctly pronounced target and do not, necessarily, suggest that hearing the mispronunciation leads to lexical retrieval of the correct pronunciation (for a similar discussion of the different implications of prelexical and lexical effects in priming see Radeau et al., 1995; Slowiaczek & Hamburger, 1992). However, the finding that a mispronunciation primes recognition of a word semantically related to the correct pronunciation – as shown in our study and in Marslen-Wilson and Zwitserlood (1989) – strongly suggests the *lexical* activation of the phonologically related canonical form upon hearing the mispronunciation.

Indeed, our results provide stronger evidence that mispronunciations activate the canonical forms of words than previous mispronunciation studies with children to date (Bailey & Plunkett, 2002; Mani & Plunkett, 2007; Swingley & Aslin, 2000; White & Morgan, 2008; among others). In these studies, children’s attention to the target object upon hearing a mispronunciation need not necessarily be interpreted as evidence for the mispronunciation activating the correct pronunciation of a word. As previous research has consistently shown, infants internally generate the labels of visually fixated familiar images (Mani & Plunkett, 2010c, 2011; Mani et al., 2012). The presentation of the target and distracter image before naming in mispronunciation detection studies allows the child to internally generate the canonical form of the target label. Attention to the target object upon hearing the mispronunciation could, therefore, be driven by children’s detection of the phonological overlap between the heard mispronunciation and the internally generated canonical form. That is, previous mispronunciation detection studies cannot conclude that hearing *gat* leads to lexical retrieval of the word *cat*, because it remains unclear whether hearing the mispronunciation would have led to activation of the correct label in the absence of visual reinforcement. Our results, in contrast, cannot be explained by a similar mechanism because children were never presented with the image of the prime and only lexical retrieval of the canonical form from the auditorally presented mispronunciation could lead to cascaded activation of the semantically related target (as only *cat*, but not *gat*, is semantically related to *dog*).

The finding that mispronunciations lead to cascaded activation of words semantically related to the canonical form even in toddlers has important implications for our understanding of the mechanisms underlying infant word recognition. First, it supports earlier suggestions that word recognition in infants as young as two years of age involves similar processes as in adults and leads to the activation of phonologically and semantically related words as well as lexical activation that cascades through different levels of processing (Arias-Trejo & Plunkett, 2009; Mani & Plunkett, 2010c; Mani et al., 2012).

Second, it strongly suggests that the infant word recognition system is robust enough to cope with imperfect input. Although the prime was mispronounced, children still showed a robust priming effect for a (phono-) semantically related target word. As stated before, this implies *lexical* activation of the canonical form of the prime and cannot be attributed to sublexical effects as only the canonical form is semantically related to the target (see also Mani et al., 2012; Marslen-Wilson & Zwitserlood, 1989). This finding expands previous findings on phono-semantic priming in toddlers (Huang & Snedeker, 2010; Mani et al., 2012) because it shows that cascaded activation need not only be initiated by familiar words that are part of the child's lexicon, but also by mispronunciations (i.e., nonwords that are not part of the child's lexicon). It also goes beyond the findings of previous dialect studies (Schmale et al., 2012; White & Aslin, 2011) in that it shows that toddlers can not only cope with systematic variation in the input, but also with unpredictable deviations (i.e., mispronunciations that do not follow a specific rule or pattern).

One could, however, argue that the children in our study did not detect the mispronunciation and therefore showed a priming effect for correctly and mispronounced primes. We suggest that this interpretation is unlikely for two reasons. First, time course analysis revealed that the difference between mispronounced trials and unrelated trials begins 200 ms later than the difference between correctly pronounced and unrelated trials, suggesting that word recognition is hindered by the mispronunciation. Although this difference in the effect of mispronounced and correctly pronounced primes appears rather late, it is consistent with the timing of the lexical interference effect reported by Von Holzen and Mani (2012) for bilingual toddlers in a cross-language auditory priming task. Second, there is no shortage of evidence that two-year-olds and even younger infants, for example, 12-month-olds (Mani & Plunkett, 2010b), are highly sensitive to subtle mispronunciations of familiar words (Bailey & Plunkett, 2002; Mani & Plunkett, 2007; Swingley & Aslin, 2000; White & Morgan, 2008; among others).

Nevertheless, it appears that there is no strong difference between the priming effects for correctly and mispronounced primes, suggesting that activation of the correct form is only mildly hindered by the mispronunciation. This leads to a third implication of our findings, namely that initial mispronunciations do not necessarily prevent word recognition. Rather, it appears that sufficient phonological overlap between a mispronunciation and its correct pronunciation can lead to lexical activation of the stored representation of a word. This questions models of word recognition that put special emphasis on the relevance of the word onset in word recognition (e.g., Marslen-Wilson, 1987) and speaks for continuous models of word recognition that consider overall phonological overlap to be more important (e.g., Luce & Pisoni, 1998; McClelland & Elman, 1986; Norris, 1994). Note that mispronunciations and correct forms overlapped considerably and only differed in the place or manner of articulation of the initial consonant. If overall phonological overlap between mispronunciation and the correct form is the crucial factor modulating lexical activation, it seems reasonable that the difference in the priming effect is marginal: The employed priming paradigm gives the children some time to process the mispronunciation before the actual target comes into play, thereby giving them time to activate the correct form based on the large phonological overlap between the acoustic signal and the stored form. Once the correct form is activated and activation cascades to semantically related words, the effects of a mispronunciation as subtle as the ones used here might result in only small differences in semantic priming.

Indeed, some previous mispronunciation studies (Swingley & Aslin, 2000, 2002; but see Ballem & Plunkett, 2005; Mani & Plunkett, 2007; White & Morgan, 2008) find that children tend to fixate an object even upon hearing a mispronunciation of the label for this object. As noted above, this could be driven by the visual context provided by the target image being presented on-screen and the greater overlap between the mispronunciation and the expected target label relative to the distracter label. It would be interesting to examine whether results similar to the current study are obtained when the (phono-) semantically related target is presented alongside a novel distracter image, thereby allowing the mispronunciation to be treated as the label for the novel image (similar to White & Morgan, 2008). Such a finding would provide an interesting contrast to the results of the current study and those of Mani et al. (2012) by showing that the presence of a novel distracter image leads to the mispronunciation being treated as a novel word, thereby blocking cascaded activation of the canonical form and its semantic associates.

Indeed, this raises the question of how the mispronunciation in the current study leads to activation of the canonical form, given that the mispronunciations were, technically, nonwords unknown to the children. In a model like TRACE (McClelland & Elman, 1986), this can be neatly captured by activation at the phoneme level leading to activation at the word level. TRACE assumes that the incoming input will activate the corresponding phonemes, that is, hearing the mispronunciation *gat* will lead to activation of the phonemes/g/,/a/, and/t/. This activation from the phoneme level can then feed through to activate lexical candidates compatible with it. As *gat* is a nonword, there will be no perfect match at the lexical level. However, the correct pronunciation *cat* (and other lexical candidates that overlap with the mispronunciation, such as *hat*) might still be activated due to the phoneme overlap between *gat* and *cat*. Similarly, a probabilistic model like Shortlist B (Norris & McQueen, 2008) would allow for recognition of the correct form upon hearing the mispronunciation. Shortlist B assumes that listeners evaluate the conditional probability of each lexical candidate given the incoming input, selecting the most probable word candidate. Thus, the correct form would be activated upon hearing the mispronunciation because it is the most likely word given the perceived input – especially given that there are hardly any other rhyming neighbors in the child's lexicon that could be better or equally fitting candidates (De Cara & Goswami, 2002). Activation of the correct form then cascades activation to words semantically related to this word, including the target word (the above-mentioned models are models of speech perception addressing phonological and not semantic processing and do not make predictions on the activation of semantically related words). Activation of these related words might be propagated through either overlap in semantic features shared by the target and the correct pronunciation (Cree & McRae, 2003; McRae & Boisvert, 1998) or through a direct link between words that have similar meaning (Collins & Loftus, 1975). Whatever the precise mechanism that leads to activation of a semantically related word, our results corroborate previous findings of (phono-) semantic priming in toddlers (Arias-Trejo & Plunkett, 2009; Mani et al., 2012) and indicate that words are organized based on phonological and semantic information in the developing lexicon.

Note that the above scenario relies on a different mechanism than the one that has been proposed for infants' ability to cope with dialectal variation. It has been argued that the adaptation to speaker idiosyncrasies can be taken as evidence for perceptual learning at a prelexical level (Cutler, Eisner, McQueen, & Norris, 2010; Samuel & Kraljic, 2009). This entails that the mechanism mapping acoustic input and phonemic representations is altered to accommodate the speaker-specific variation (see also White & Aslin, 2011). Thus, the deviant dialectal pronunciations would get 'normalized' to map the standard pronunciation *before* lexical access takes place. In case of unpredicted mispronunciations no such normalization can occur. Instead, toddlers might map the deviant pronunciation to the standard pronunciation by detecting the phonological overlap between the two *while* lexical access takes place (as described above). Thus, the way in which lexical activation of the canonical form is achieved differs across the two scenarios. Yet, the mechanism that underlies the *initial* perception of accented words (i.e., the first encounter with accented words before speaker or dialect accommodation takes place) might be very similar to the mechanism described for mispronounced words here: after all, the child has to access a word based on imperfect input in both cases. Regardless of the differences between the scenarios, our results show that subtle mispronunciations lead to activation of the correct form and cascades to words lexically related to this correct form. This speaks to the flexibility of the word recognition system and provides a crucial step in investigating how word recognition develops in infants.

ACKNOWLEDGMENTS

This research was supported by the German Excellence Initiative award to Georg-August-University Göttingen (Third funding line: Institutional strategy).

REFERENCES

Altvater-Mackensen, N. (2010). Do manners matter? Asymmetries in the acquisition of manner of articulation features. Doctoral Dissertation, Radboud University Nijmegen.

Altvater-Mackensen, N., & Mani, N. (under review). Do phonological features mediate lexical access in infants?

Arias-Trejo, N., & Plunkett, K. (2009). Lexical-semantic priming effects in infancy. *Philosophical Transactions of the Royal Society of London: Series B, Biological Sciences*, *364*, 3633–3647.

Bailey, T. M., & Plunkett, K. (2002). Phonological specificity in early words. *Cognitive Development*, *17*, 1267–1284.

Ballem, K. D., & Plunkett, K. (2005). Phonological specificity in children at 1;2. *Journal of Child Language*, *32*, 159–173.

Best, C. T., Tyler, M. D., Gooding, T. N., Orlando, C. B., & Quann, C. A. (2009). Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words. *Psychological Science*, *20*, 539–542.

Canfield, R. L., Smith, G. E., Brezsnyak, M. P., Snow, K. L., Aslin, R. N., Haith, M. M., . . . Adler, S.A. (1997). Information processing through the first year of life: A longitudinal study using the visual expectation paradigm. *Monographs of the Society for Research in Child Development*, *62*(2), Serial No. 250.

Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review*, *82*, 407–428.

Connine, C. M., Blasko, D. G., & Titone, D. (1993). Do the beginnings of spoken words have a special status in auditory word recognition? *Journal of Memory and Language*, *32*, 193–210.

Cree, G. S., & McRae, K. (2003). Analyzing the factors underlying the structure and the meaning of chipmunk, cherry, chisel, cheese and cello (and many other such concrete nouns). *Journal of Experimental Psychology: General*, *132*, 163–201.

Creel, S. C. (2012). Phonological similarity and mutual exclusivity: On-line recognition of atypical pronunciations in 3–5-year-olds. *Developmental Science*, *15*, 697–713.

Cutler, A., Eisner, F., McQueen, J. M., & Norris, D. (2010). How abstract phonemic categories are necessary for coping with speaker-related variation. In C. Fougeron, B. Kühnert, M. D'Imperio, & N. Vallée (Eds.), *Laboratory Phonology*. Vol. 10 (pp. 91–111). Berlin: Mouton de Gruyter.

De Cara, B., & Goswami, U. (2002). Statistical analysis of similarity relations among spoken words: Evidence for the special status of rimes in English. *Behavioural Research Methods and Instrumentation*, *34*, 416–423.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, *59*(5), Serial No. 242.

Havy, M., & Nazzi, T. (2009). Bias for consonantal over vocalic information in word learning at 16 months. *Infancy*, *14*, 439–456.

Huang, Y. T., & Snedeker, J. (2010). Cascading activation across levels of representation in children's lexical processing. *Journal of Child Language*, *38*, 644–661.

Huettig, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, *57*, 460–482.

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

Lukatela, G., Eaton, T., Lee, C., & Turvey, M. T. (2001). Does visual word identification involve a sub-phonemic level? *Cognition*, *78*, B41–B52.

Mani, N., Coleman, J., & Plunkett, K. (2008). Phonological specificity of vocalic features at 18 months. *Language and Speech*, *51*, 3–21.

Mani, N., Durrant, S., & Floccia, C. (2012). Activation of phonological and semantic codes in toddlers. *Journal of Memory and Language*, *66*, 612–622.

Mani, N., & Plunkett, K. (2007). Phonological specificity of consonants and vowels in early lexical representations. *Journal of Memory and Language*, *57*, 252–272.

Mani, N., & Plunkett, K. (2010a). Does size matter? Graded sensitivity to vowel mispronunciations of familiar words. *Journal of Child Language*, *38*, 606–627.

Mani, N., & Plunkett, K. (2010b). Twelve-month-olds know their cups from their keps and tups. *Infancy*, *15*, 445–470.

Mani, N., & Plunkett, K. (2010c). In the infant's mind's ear: Evidence for implicit naming. *Psychological Science*, *21*, 908–913.

Mani, N., & Plunkett, K. (2011). Phonological priming and cohort effects in toddlers. *Cognition*, *121*, 196–206.

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*, 177–190.

Marslen-Wilson, W. D. (1987). Functional parallelism in spoken word recognition. *Cognition*, *25*, 71–102.

Marslen-Wilson, W., & Zwitserlood, P. (1989). Accessing spoken words: The importance of word onsets. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 576–585.

McClelland, J. L., & Elman, J. L. (1986). The TRACE model of speech perception. *Cognitive Psychology*, *18*, 1–86.

McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 863–883.

Meints, K., & Woodford, A. (2008). *Lincoln Infant Lab Package 1.0: A new programme package for IPL, Preferential Listening, Habituation and Eyetracking*. [WWW document: Computer software & manual]. URL: http://www.lincoln.ac.uk/psychology/babylab.htm

Meyer, A. S., Belke, E., Telling, A., & Humphreys, G. W. (2007). Early activation of object names in visual search. *Psychonomic Bulletin & Review*, *14*, 710–716.

Milberg, W., Blumstein, S. E., & Dworetzky, B. (1988). Phonological factors in lexical access: Evidence from an auditory lexical decision task. *Bulletin of the Psychonomic Society*, *26*, 305–308.

Misiurski, C., Blumstein, S. E., Rissman, J., & Berman, D. (2005). The role of lexical competition and acoustic-phonetic structure in lexical processing: Evidence from normal subjects and aphasic patients. *Brain and Language*, *93*, 64–78.

Nazzi, T. (2005). Use of phonetic specificity during the acquisition of new words: Differences between consonants and vowels. *Cognition*, *98*, 13–30.

Nazzi, T., Floccia, C., Moquet, B., & Butler, J. (2009). Bias for consonantal over vocalic information in French- and English-learning 30-month-olds: Crosslinguistic evidence in early word learning. *Journal of Experimental Child Psychology*, *102*, 522–537.

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition*, *52*, 189–234.

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*, 357–395.

Radeau, M., Morais, J., & Segui, J. (1995). Phonological priming between monosyllabic spoken words. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 1297–1311.

Samuel, A. G., & Kraljic, T. (2009). Perceptual learning for speech. *Attention, Perception and Psychophysics*, *71*, 1207–1218.

Schmale, R., Christia, A., & Seidl, A. (2012). Toddlers recognize words in an unfamiliar accent after brief exposure. *Developmental Science*, *15*, 732–738.

Schmale, R., Hollich, G., & Seidl, A. (2011). Contending with foreign accent in early word learning. *Journal of Child Language*, *38*, 1096–1108.

Slowiaczek, L. M., & Hamburger, M. B. (1992). Prelexical facilitation and lexical interference in auditory word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1239–1250.

Snow, C. E., & Ferguson, C. A. (Eds.) (1977). *Talking to children: Language input and acquisition*. New York: C.U.P.

Stokes, S. (2010). Neigborhood density and word frequency predict vocabulary size in toddlers. *Journal of Speech, Language, and Hearing Research*, *53*, 670–683.

Swingley, D. (2009). Onset and codas in 1.5-year-olds' word recognition. *Journal of Memory and Language*, *60*, 252–269.

Swingley, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, *76*, 147–166.

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*, 480–484.

Szagun, G., Stumper, B., & Schramm, S. A. (2009). *FRAKIS Fragebogen zur frühkindlichen Sprachentwicklung*. Frankfurt: Pearson.

Van der Feest, S. V. H. (2007). Building a phonological lexicon. The acquisition of the Dutch voicing contrast in perception and production. Doctoral Dissertation, Radboud University Nijmegen.

Van Heugten, M., & Johnson, E. K. (2013). Learning to contend with accents in infancy: Benefits of brief speaker exposure. *Journal of Experimental Psychology: General*. doi: 10.1037/a0032192.

Von Holzen, K., & Mani, N. (2012). Language non-selective lexical access in bilingual toddlers. *Journal of Experimental Child Psychology*, *113*, 569–586.

White, K. S., & Aslin, R. N. (2011). Adaptation to novel accents by toddlers. *Developmental Science*, *14*, 372–384.

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*, 114–132.

Yee, E., & Sedivy, J. C. (2006). Eye movements to pictures reveal transient semantic activation during spoken word recognition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 1–14.

Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, *125*, 244–262.