Going with TRACE beyond Infant Mispronunciation Studies: Lexical Networks and Phoneme Competition

Julien Mayor (julien.mayor@unige.ch)
FPSE, University of Geneva
1211 Genève 4, Switzerland

Kim Plunkett (kim.plunkett@psy.ox.ac.uk)
Department of Experimental Psychology, University of Oxford
Oxford OX1 3UD, United Kingdom

Abstract

The TRACE model of speech perception (McClelland & Elman, 1986) is used to simulate graded sensitivity to mispronunciations of familiar words as reported by White and Morgan (2008). Our simulations predict that phoneme or lexical competition may be absent in the mental lexicons of the 19-month-old infants tested experimentally.

Keywords: Word learning; speech perception; language acquisition; inhibition

Introduction

Research on infant spoken word recognition has made dramatic advances over the past two decades. Spurred on by the refinement of experimental techniques such as the familiarisation head turn preference procedure (Jusczyk & Aslin, 1995), the switch task (Stager & Werker, 1997) and the mispronunciation task (Swingley & Aslin, 2000), our understanding of *what* infants and young children know about the sounds of words, both familiar and newly learnt, has expanded incrementally. However, our appreciation of the representations and processes underlying early phono-lexical knowledge and *how* these develop is less advanced. Although these approaches offer important insights as to how infants and young children develop knowledge about the sounds of words, they do not provide a precise computational account of the representations and processes involved. In this paper, we describe our attempt to apply the TRACE model of word recognition (McClelland & Elman, 1986) to simulate aspects of spoken word recognition during infancy and early childhood.

TRACE was originally proposed as a model of adult spoken word recognition. In TRACE, spoken word recognition is modelled as an incremental process involving the elimination of competing candidates that are represented in the individual's mental lexicon. Various accounts have emphasised the role of cohort competitors (Marlsen-Wilson & Welsh, 1978) and phonological neighbours (Cutler, 1995; Goldinger, Luce, & Pisoni, 1989) in this competition. Allopenna, Magnuson, and Tanenhaus (1998) have argued that the TRACE model of speech perception provides a satisfactory accommodation for the role of cohorts and phonological neighbours in the resolution of the competitive process.

In TRACE, acoustic-phonetic features are mapped over time onto phoneme nodes that map onto lexical nodes, with lexical-phonemic feedback and lateral inhibition at the phonemic and lexical levels (see Fig. 1).

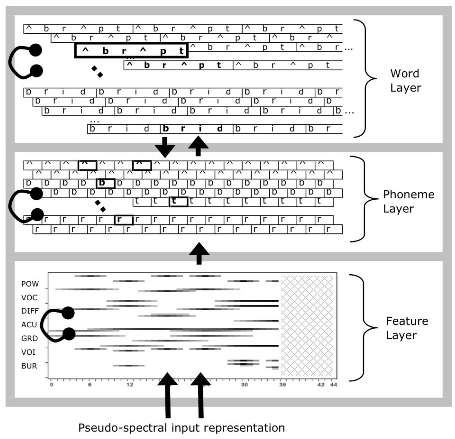

Figure 1: Schematic diagram of TRACE architecture. Drawing by Ted Strauss.

Allopenna et al. (1998) found that the time course of spoken word recognition indexed by eye movements in human participants can be modelled by such continuous mapping models: Adults were instructed to move one of four objects that were on a screen, while they were simultaneously monitored by an eye-tracker. Along with the referent, three competitors were displayed on screen; a cohort competitor (object starting with the same onset and vowel), a rhyme competitor and an unrelated competitor. Using the TRACE model, implementing a forced choice with Luce's choice rule (Luce, 1959), Allopenna et al. (1998) accurately reproduced the typical pattern of eye-gaze of the participants. For example, adults were likely to be distracted by both cohort and rhyme competitors in this task. TRACE also exhibits enhanced activation for these competitors resulting in enhanced levels of "eye fixations" when using Luce's choice rule. More recently, TRACE has been used to model adult's gradient sensitivity to within-category voice onset time manipulations in a visual world task (McMurray, Tanenhaus, & Aslin, 2009) and individual differences in online spoken word recognition, including individuals at risk for specific language impairment (McMurray, Samelson, Lee, & Bruce Tomblin, 2010). In both these applications, exploration of TRACE's parameter space identified factors (phoneme inhibition and lexical decay, respectively) that might account for observed human performance.

We adopt the same approach and try and refine the potential architecture underlying infant word recognition by simulating with TRACE the finding that infants display a graded sensitivity to the severity of mispronunciations (White & Morgan, 2008).

Graded sensitivity to the severity of mispronunciations; Implications

Infants show graded sensitivity to mispronunciations of familiar words, as a function of the severity of the mispronunciation. White and Morgan (2008) have shown that 19-month-olds show a graded response in their looking behaviour when presented, in an Inter-Modal Preferential Looking (IPL, Golinkoff, Hirsh-Pasek, Cauley, & Gordon, 1987) task, with a correct pronunciation, 1-feature, 2-feature or 3-feature mispronunciation of the onset consonant of a target word: Infants look longer at the target object when supplied with more accurate renditions of the target object's name. In their experiment, the two pictures corresponded to a target object and a novel object. In contrast to other mispronunciation experiments (Mani & Plunkett, 2007; Swingley & Aslin, 2000), the distracter image is name-unknown and thus does not represent a potential competing lexical entry as it is a novel image. White and Morgan (2008) argued that using a novel object as a distracter is important for demonstrating graded sensitivity as it offers the infant the opportunity to consider the mispronunciation as a label for the novel distracter. This possibility is not available to the infant when the distracter is a name-known object. On the basis of their experimental findings, White and Morgan (2008) argue that lexical processing in toddlers is affected by sub-segmental phonological detail. In this simulation, we examine the adaptations of the TRACE architecture that are needed to simulate the (White & Morgan, 2008) results, and explore the ramifications of these adaptations for interpreting their experimental findings.

Method

We used jTRACE (Strauss, Harris, & Magnuson, 2007), a re-implementation of the TRACE model (McClelland & Elman, 1986). We created typical lexicons for 18 month olds by compiling British CDIs ((Hamilton, Plunkett, & Schafer, 2000), a British adaptation of the MacArthur-Bates CDI, (Fenson et al., 1993)) using words that are understood by at least 50% of the infants at 18 months of age. The lexicon is specified using data from 179 infants and count 131 words.

Recognition time for spoken words is affected not only by the number of phonological neighbours (Cutler, 1995), but also by their frequency (Goldinger et al., 1989). Therefore, we identified individual token frequencies, by extracting word frequencies on all tiers based on the Manchester corpora (Theakston, Lieven, Pine, & Rowland, 2001) from the CHILDES database (MacWhinney, 1991), where 12 English children were recorded weekly from 20 to 36 months of age. Word frequencies used in the simulations are raw word counts on the whole corpora, converted to frequency per million. When implementing frequencies in the model, we follow the suggestions advocated by MacKay (1982) and implemented by Dahan, Magnuson, and Tanenhaus (2001), i.e., frequency modulates the connection weights associated with lexical units, using the same value for the scaling parameter (0.13) used in (Dahan et al., 2001). The modulation of frequency effects via phoneme-lexicon connection strengths is consistent with a learning basis for frequency (e.g., of the Hebbian type). In addition, Dahan et al. (2001) found this type of bottom-up connection strength implementation to have qualitative advantages over resting state and post-perceptual frequency manipulations.

Given the large size of the infant lexicon at 18 months of age, many of the phonemes needed to represent the different words were not encoded in the original TRACE model (McClelland & Elman, 1986) nor in its re-implementation (Strauss et al., 2007). Therefore, we added feature values for all phonemes used in the infant's lexicon[1].

Correctly pronounced words and mispronounced words are presented to the model and activation levels of two competitors (the target and a distracter) are monitored. We adopt the same linking hypothesis as (Allopenna et al., 1998) in order to map the activation levels to fixation durations. Activation of a word is the result of both its direct activation due to phonological overlap with the input and the result of competition with all other words that are activated with that same input. Only items that are on display are available as potential responses. Similarly to (Allopenna et al., 1998), the activation levels a of the displayed items are then transformed into response strengths following (Luce, 1959). Given the high salience of the images, we assume that total looking time is split entirely between the target and distractor objects, enabling us to convert the response strengths into fixation durations using the Luce choice rule. The proportion of looking to the target at time t is given by: $p_{target}(t) = \frac{e^{ka_{target}(t)}}{e^{ka_{target}(t)} + e^{ka_{distractor}(t)}}$ where k is a free parameter determining the amount of separation between units of different activations (value set to $k = 2$). All other parameters used in jTRACE were set to their default values. Proportion of looking times to the target and distracters are reported as the average over 100 processing cycles starting with the onset of the pronounced word.

We used the stimuli described in Experiment 1 of (White & Morgan, 2008), reproduced in Table 1, with the exception of the word "cookie", which is not present in the British version of the CDI that we used to create the new jTRACE dictionaries. Since the distractor is name-unknown in the White and Morgan (2008) experiment, the activation level associated with the novel object on display is set to zero. It is noteworthy that, however, due to the application of Luce's

[1]Thanks to Ōiwi Parker-Jones for help in assigning feature values for phonemes not present in the original TRACE model.

rule, both images share some amount of the total looking time spent during each trial[2].

Table 1: Correctly pronounced and mispronounced labels presented to infants in Experiment 1 of White & Morgan (2008). The unfamiliar words used by White and Morgan (2008) are not listed here because they do not compete for recognition in TRACE. The table also includes the cohort size as a function of pronunciation type for the stimuli used in White & Morgan (2008).

Correct pronunciations	Mispronunciations 1-feature	2-feature	3-feature
keys	teys	deys	zeys
book	dook	took	sook
bear	gear	tear	sear
foot	soot	zoot	goot
car	par	dar	zar
ball	gall	kall	sall
bird	gird	kird	sird
bottle	gottle	kottle	sottle
shoe	foe	voe	goe
cup	tup	bup	vup
hand	fand	zand	dand
Mean cohort size (SD)			
18.7 (12.1)	7.7 (7.2)	11.7 (9.9)	4.4 (2.5)

In this approach, mispronunciations cannot act as potential labels for the distracter image since the distracter image is name-unknown. The unfamiliar words used by White and Morgan (2008) do not belong to the lexicon, and therefore do not compete for recognition in TRACE. Simulations were run with the 18-month-lexicon to mimic the behavior of 19-month olds.

Results

First, we ran simulations with jTRACE's default parameters for the same stimuli used by White and Morgan (2008). The top panel of Figure 2 depicts the proportion of looking time associated with the target in the correct, 1-feature-, 2-feature- and 3-feature-mispronunciations. No graded sensitivity is observed as a function of the severity of mispronunciation. Since the metrics used by White and Morgan (2008) to derive the severity of mispronunciation may differ slightly form jTRACE's, we also evaluate the impact of the severity of mispronunciations on the level of activation of the target words within jTRACE's metrics. The bottom panel of Figure 2 depicts the reduction in activation level as a function of the magnitude of the mispronunciation (Euclidean distance between the two phonemes in jTRACE's feature space) for all stimuli. The absence of any correlation suggests that activation levels of target words are not directly sensitive to the severity of mispronunciations, in contrast to White and Morgan (2008) findings.

Closer examination of the stimuli used by White and Morgan (2008) reveals that the number of cohort competitors in the typical lexicon of an 18-month old differs dramatically with mispronunciation type. Table 1 presents an analysis

[2] $p_{target}(t)$ being an exponential function of word activation.

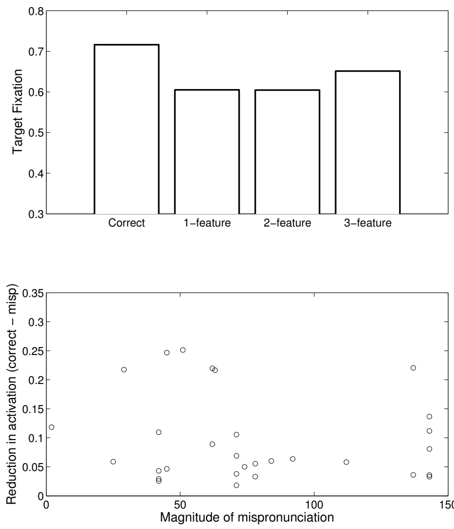

Figure 2: Top Panel: Simulation of White and Morgan (2008) with jTRACE's default parameters. The unbalanced cohort sizes in each condition interferes with the bottom-up activation flow favoring graded sensitivity to the severity of the mispronunciations. In particular, looking times in the three-feature mispronunciation condition are longer than in the one- and two-feature mispronunciation conditions. Bottom panel: Mispronunciation effect (reduction in activation due to the mispronunciation) as a function of the magnitude of the mispronunciation in jTRACE's feature space. No correlation is observed between looking times and the severity of mispronunciations.

of the cohort size associated with correct pronunciations and each mispronunciation type. It is apparent that 3-feature mispronunciations have far fewer cohort competitors than any of the other mispronunciation conditions. An item-analysis of variance of the number of cohort competitors across types of pronunciation yielded a main effect of pronunciation condition (F=5.53, df=3, p=0.0028). Two feature mispronunciations have marginally more cohort competitors than 1-feature mispronunciations (t=1.34, df=10, p=0.21, n.s.), and more importantly, more than 3-feature mispronunciations (t=2.40, df=10, p=0.038).

An important characteristic of TRACE is that it implements competition within the different layers of the network. As a consequence, cohort competitors impact the activation levels associated with a target word. A low number of cohort competitors leads to reduced inhibition which, in turn, leads to higher activation of the target word. For the stimuli used in (White & Morgan, 2008), we expect the cohort competition in TRACE to interfere with any mispronunciation effect. In particular, the low number of cohort competitors in the case of the 3-feature mispronunciation would lead to an *increase* in the activation of the target word, rather than to a *decrease* in its activation level. Clearly, this outcome would be incommensurate with White & Morgan's finding of a graded sensitivity to severity in the mispronunciation and explains why a graded sensitivity to the severity of mispronunciations was not observed with jTRACE's default parameters. Therefore, we conducted a series of simulations so as to evaluate the impact of word-layer and phoneme-layer inhibition on sensitivity to mispronunciation.

First, we investigate the impact of reducing the level of lexical inhibition. Both theoretical and experimental considerations motivate this adaptation of TRACE: Lexical inhibition may be reduced in infancy due to the sparseness of the lexical space (Gaskell & Marslen-Wilson, 1997). Also, several recent experimental findings provide evidence that word to word interactions do not reach adult levels of competition before about 21 months of age. For example, Arias-Trejo and Plunkett (2009) and Styles and Plunkett (2009) used a semantic priming task with infants to demonstrate evidence for lexico-semantic networks in 21- and 24-month old infants. However, they failed to find evidence of semantic priming in 18-month olds. Arias-Trejo and Plunkett (2009) suggest that entries in the 18-month old lexicon may be best characterised in terms of *lexical islands* that are not in competition with each other because they are unconnected. More direct evidence is provided in a phonological priming task (Mani & Plunkett, 2011) conducted with 18- and 24-month old infants. Mani and Plunkett (2011) reported cohort effects in 24-month olds (less target looking for words from large cohorts than words from small cohorts) but no cohort effects for 18-month olds. It is likely that these age differences in cohort effects are driven by differences in the vocabulary sizes of the infants involved in the study, even though both age groups were tested on the same set of words. This set of findings, together with the findings from (Arias-Trejo & Plunkett, 2009) and (Styles & Plunkett, 2009), provide a convergent rationale for reducing lexical competition in the simulation of White & Morgan's 19-month old infants.

The top panel of Figure 3 displays the proportion of target looking in jTRACE associated with the stimuli used by White and Morgan (2008) for correct, 1-feature, 2-feature and 3-feature mispronunciations when lexical inhibition is essentially turned off ($C = 0.0001$)[3]. A graded sensitivity to the severity of mispronunciations emerges, similar to the 19-month-olds tested by White and Morgan (2008). However, correlations between the reduction of activation levels associated with target words and the magnitude of the mispronunciations in TRACE's feature space did not reach significance ($p = 0.13$, see bottom panel of Figure 3). For $C \geq 0.001$, cohort effects counteract the effect of mispronunciation such that the activity level associated with the 3-feature mispronunciations is higher that the activity level associated with the 2-feature mispronunciations.

[3]For comparison, the value commonly used to model adult sensitivities to mispronunciations is $C = 0.03$ (see for example (Allopenna et al., 1998)) which means inhibition in the word layer is 300 times stronger than the value used here.

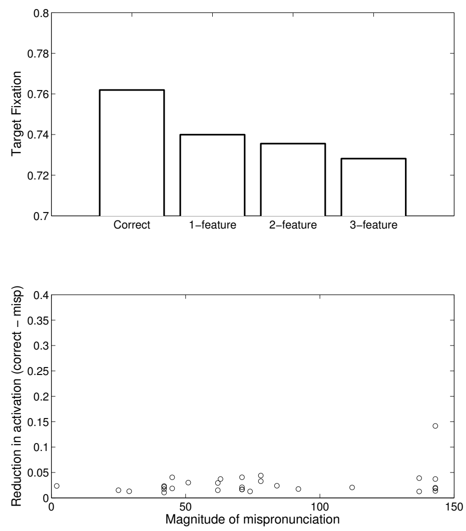

Figure 3: Top Panel. Simulation of White and Morgan (2008) with TRACE with reduced lexicon competition. Cohort effects are reduced and the bottom-up activation flow favoring graded sensitivity to the severity of the mispronunciations is not disrupted. Bottom panel. Mispronunciation effect (reduction in activation due to the mispronunciation) as a function of the magnitude of the mispronunciation in TRACE's feature space. A weak, non-significant, correlation is observed between looking times and the severity of mispronunciations.

A second manipulation that may lead to a reduction in the influence of imbalanced cohort sizes when simulating White & Morgan's findings is to reduce phoneme inhibition. McMurray et al. (2009) suggest that phoneme-level inhibition in TRACE is incompatible with recovery from "lexical garden-paths" initiated by ambiguous phonemes early in a word. We now consider the impact that the absence of phoneme-level inhibition may have on simulations of White & Morgan's findings. The top panel of Figure 4 depicts the proportion of looking time at the target when correctly pronounced, and with three levels of mispronunciation severity, when phoneme level inhibition is eliminated in TRACE. A clear, graded reduction in activation level emerges as the number of feature changes increases. Furthermore, the bottom panel of Figure 4 indicates that, within TRACE's feature metrics, a significant correlation ($R = 0.753$, $p = 1.56 \cdot 10^{-6}$) is present between the magnitude of the mispronunciation and its impact on activation levels. Cohort effects are effectively reduced and the bottom-up flow from the feature level to the lexical level, via the phoneme level, is not disrupted by cohort effects.

Discussion

(White & Morgan, 2008) reported a graded sensitivity in 19-month old infants to the severity of the mispronunciation of

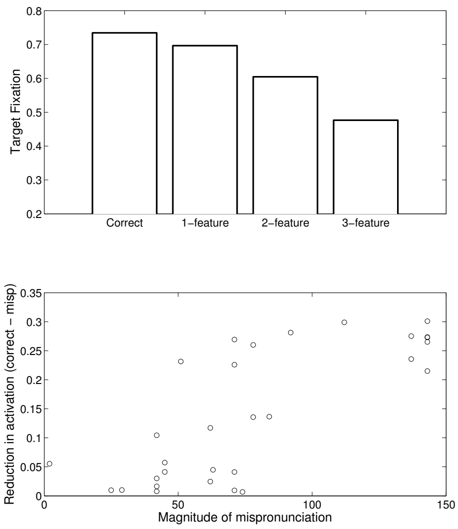

Figure 4: Top Panel: Simulation of White and Morgan's (2008) findings in jTRACE with no phoneme inhibition. Cohort effects are reduced and the bottom-up activation flow favouring graded sensitivity to the severity of the mispronunciations is well established. Bottom panel: Mispronunciation effect (reduction in activation due to the mispronunciation) as a function of the magnitude of the mispronunciation in TRACE's feature space. A strong, significant, correlation is observed between target preference and the severity of mispronunciations.

a target word and argued that this finding demonstrated fine-grained sensitivity at the sub-segmental level. The gradual decrease in looking time at the target object as the number of modified features increased was observed despite the fact that the number of cohort competitors for mispronunciations, as evaluated by an analysis of CDI reports, was smaller for the 3-feature mispronunciations than for the 2-feature mispronunciations. Competition between word activation levels in TRACE has an opposite effect on target word activation to mispronunciation severity for the stimuli used by White and Morgan (2008), leading to an apparent incompatibility between White & Morgan's findings and the predictions of jTRACE . The fact that White and Morgan (2008) report that target looking decreased with mispronunciation severity suggests that either inhibition between competing words in the lexicon is not present (or extremely reduced) at 19 months of age (consistent with Mani and Plunkett (2011); Arias-Trejo and Plunkett (2009) or that phoneme-level inhibition should be removed (consistent with McMurray et al., 2009).

An alternative possibility is that the apparent asymmetry between cohort sizes used in (White & Morgan, 2008) is illusory. It is recognised that parental reports provide under-estimates of actual vocabulary sizes (Mayor & Plunkett, 2011). A proper estimate of vocabulary composition may result in a more balanced lexicon structure, in turn reducing the impact of cohort imbalance disrupting the graded sensitivity to mispronunciation severity. However, an analysis of a dense recording at 30 months of age, the Haggerty corpus (Haggerty, 1929), revealed that /b/-onset words (89 words) are almost twice as numerous as /p/-onset words (48 words). Better descriptions of the lexical composition in infancy would no doubt help refine the distribution of cohort sizes associated with different onsets. However, they are unlikely to reveal an even profile in cohort sizes.

Taken individually, neither a reduction in lexical-level inhibition, the removal of phoneme-level inhibition, nor a finer-grained estimate of vocabulary composition in infancy can fully account for the graded sensitivity to mispronunciations described in (White & Morgan, 2008) while also capturing the findings that both onset consonant and medial-vowel mispronunciations lead to a reduction in target preferences reported by Mani and Plunkett (2007)[4]. A proper explanation of both phenomena will likely incorporate all of these explanations to a certain degree.

In an attempt to adjudicate between these different hypotheses, or to confirm the contribution of multiple contributing factors (reduction in overall inhibition and a slightly more balanced lexicon), one might ask whether 24-month olds would also display graded sensitivity to the severity of mispronunciations. Indeed, an important prediction of the TRACE simulation of White & Morgan's (2008) results is that graded sensitivity to mispronunciation severity will be affected by cohort and neighbourhood effects if lexical competition is active. We justified switching off lexical competition in the model on the grounds that empirical studies have reported lexical island effects and lack of cohort effects with 18-month old infants (Arias-Trejo & Plunkett, 2009; Mani & Plunkett, 2011). However, these studies also report that lexical competition effects are apparent by 24-months of age. If the lexical-level inhibition hypothesis holds, we would predict, therefore, that when a task like White & Morgan's (2008) study is conducted with 24-month-old infants, then the impact of severity of mispronunciation is likely to diminish when using the same stimuli.

It is noteworthy that the acceleration of rapid word learning, often dubbed the "vocabulary spurt" (Bloom, 1973), between 18 and 21 months of age coincides with the potential emergence of lexical competition. Of course, TRACE only implements lexical competition at a phonological level. Lexico-semantic competition, which is outside the purview of TRACE, may follow a different developmental trajectory and lead to different patterns of competition.

Finally, it should be noted that many simplifying assumptions were adopted in the simulations reported in this research. The dictionaries used in the simulations were created by assessing typical vocabularies as assessed by the Oxford CDI (Hamilton et al., 2000). However, individual differences in lexicon sizes and composition would lead to a

[4] A full analysis in reported in (Mayor & Plunkett, *Submitted*).

distribution of phonological sensitivities and looking patterns rather than a single uniform result in TRACE for a given age group. Moreover, the nonlinear impact of lexical competition in TRACE implies that a mean looking pattern based on a mean lexicon would not match the mean of looking patterns associated with different lexicon sizes. Fitting TRACE to individual lexicons rather than a standardised lexicon would provide yet another series of novel experimental predictions against which to evaluate the model.

Acknowledgments

This work is supported by the Swiss National Science Foundation grant 131700 awarded to Julien Mayor and by the Economic and Social Research Council Grant RES-062-23-0194 awarded to Kim Plunkett.

References

Allopenna, P., Magnuson, J., & Tanenhaus, M. (1998). Tracking the Time Course of Spoken Word Recognition Using Eye Movements: Evidence for Continuous Mapping Models. *Journal of Memory and Language*, *38*(4), 419–439.

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

Bloom, L. (1973). *One word at a time: The use of single word utterances*. The Hague: Mouton.

Cutler, E. (1995). *Spoken-word recognition.* San Diego: Academic Press.

Dahan, D., Magnuson, J., & Tanenhaus, M. (2001). Time course of frequency effects in spoken-word recognition: Evidence from eye movements. *Cognitive Psychology*, *42*(4), 317–367.

Fenson, L., Dale, P., Reznick, S., Thal, D., Bates, E., Hartung, J., et al. (1993). *Macarthur communicative development inventories: User's guide and technical manual*. San Diego: Singular Press.

Gaskell, M., & Marslen-Wilson, W. (1997). Integrating form and meaning: A distributed model of speech perception. *Language and cognitive Processes*, *12*(5-6), 613–656.

Goldinger, S., Luce, P., & Pisoni, D. (1989). Priming lexical neighbors of spoken words: Effects of competition and inhibition. *Journal of Memory and Language*, *28*(5), 501–518.

Golinkoff, R., Hirsh-Pasek, K., Cauley, K., & Gordon, L. (1987). The eyes have it: lexical and syntactic comprehension in a new paradigm. *Journal of Child Language*, *14*, 23-46.

Haggerty, L. (1929). What a two-and-one-half-year-old child said in one day. *Journal of Genetic Psychology*, *38*, 75-100.

Hamilton, A., Plunkett, K., & Schafer, G. (2000). Infant vocabulary development assessed with a British communicative development inventory. *Journal of Child Language*, *27*, 689–705.

Jusczyk, P., & Aslin, R. N. (1995). Infant's detection of sound patterns of words in fluent speech. *Cognitive Psychology*, *29*, 1–23.

Luce, R. (1959). *Individual choice behavior*. Wiley New York.

MacKay, D. (1982). The problems of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. *Psychological Review*, *89*(5), 483–506.

MacWhinney, B. (1991). *The CHILDES project : Tools for analyzing talk*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, *57*(2), 252–272.

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

Marlsen-Wilson, W., & Welsh, A. (1978). Processing interactions and lexical access dunng words recognition in contmuous speech. *Cognitive Psychology*, *10*, 29–63.

Mayor, J., & Plunkett, K. (2011). A statistical estimate of infant and toddler vocabulary size from cdi analysis. *Developmental Science*, *14*(4), 769–785.

Mayor, J., & Plunkett, K. (*Submitted*). *Infant Word Recognition: Insights from TRACE Simulations.*

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

McMurray, B., Samelson, V., Lee, S., & Bruce Tomblin, J. (2010). Individual differences in online spoken word recognition: Implications for sli. *Cognitive psychology*, *60*(1), 1–39.

McMurray, B., Tanenhaus, M., & Aslin, R. (2009). Within-category vot affects recovery from. *Journal of memory and language*, *60*(1), 65–91.

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than word learning tasks. *Nature*, *388*, 381–382.

Strauss, T., Harris, H., & Magnuson, J. (2007). jTRACE: A reimplementation and extension of the TRACE model of speech perception and spoken word recognition. *Behavior Research Methods*, *39*(1), 19.

Styles, S., & Plunkett, K. (2009). How do infants build a semantic system? *Language and Cognition*, *1*, 1–24.

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

Theakston, A., Lieven, E., Pine, J., & Rowland, C. (2001). The role of performance limitations in the acquisition of verb-argument structure: an alternative account. *Journal of Child Language*, *28*(01), 127–152.

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