University of Iowa
Iowa Research Online

Spring 2013

Not everything that competes means something: evidence for competition among word-forms in a novel-word learning paradigm

Efthymia Evangelia Kapnoula
University of Iowa

Recommended Citation

Kapnoula, Efthymia Evangelia. "Not everything that competes means something: evidence for competition among word-forms in a novel-word learning paradigm." MA (Master of Arts) thesis, University of Iowa, 2013.
https://doi.org/10.17077/etd.lockz7ug

NOT EVERYTHING THAT COMPETES MEANS SOMETHING:
EVIDENCE FOR COMPETITION AMONG WORD-FORMS
IN A NOVEL-WORD LEARNING PARADIGM

by

Efthymia Evangelia Kapnoula

A thesis submitted in partial fulfillment of the
requirements for the Master
of Arts degree in Psychology
in the Graduate College of
The University of Iowa

May 2013

Thesis Supervisor: Associate Professor Bob McMurray

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL
_____

MASTER'S THESIS
_____

This is to certify that the Master's thesis of

Efthymia Evangelia Kapnoula

has been approved by the Examining Committee for the thesis requirement for the Master of Arts degree in Psychology at the May 2013 graduation.

Thesis Committee: _____
                            Bob McMurray, Thesis Supervisor

                      _____
                            Prahlad Gupta

                      _____
                            Susan Wagner Cook

στην Πιπίτσα

ACKNOWLEDGMENTS

ABSTRACT

The aim of the present study was to investigate whether learning a meaningless phonological word-form, can affect its ability to compete with other words shortly after it was learned. According to previous experimental work we expected that a semantic referent (Leach & Samuel, 2007), and/or consolidation over a significant amount of time (Gaskell & Dumay, 2003) are necessary for a novel word-form to be able to engage in lateral inhibition with other words. In order to examine this we used the experimental design that was used by Dahan, Magnuson, Tanenhaus and Hogan (2001). Experiment 1 was a replication of the Dahan et al (2001) study. In Experiment 2 we added a condition in which a novel word was now assigned the role of the competitor, by inserting a nonword learning task (that was performed right before the Dahan task). The goal was to see whether any differences would arise between this new novel-word condition and the nonword condition. The results from Experiment 2 were inconclusive due to the stimulus set and this is why we conducted Experiment 3, which was similar to Experiment 2, but had a different stimulus set. The results of Experiment 3 showed that, in contrast to the predictions, a novel word can compete with other words, even if it does not have meaning and, moreover, this happens immediately after training. These findings indicate that 1) a word does not have to be complete (i.e. include semantic information) in order to compete with other words and 2) connections between novel and known words can form faster than what has been suggested.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

To successfully learn a word, learners must acquire a mental representation that includes information about its phonological properties (the word-form), its meaning and the link between them (Gupta & Tisdale, 2009). However, the lexicon is no longer seen as a mental dictionary, a list of facts and information about each word (c.f., Elman, 2009). Rather, words play a pivotal role in processing, playing an organizing role in syntactic parsing (Altmann & Kamide, 1999; MacDonald, Pearlmutter & Seidenberg, 1994; Tanenhaus & Trueswell, 1995), and in organizing speech perception (Ganong, 1980; McClelland, Mirman & Holt, 2006; Luce & Pisoni, 1998). That is, words in the lexicon are not just a static collection of information. Rather, this information is stored in the links among words, and in the links between words and other representations (e.g., between words and phonemes). In such an architecture, real-time interactions between mental representations of words are fundamental to cognitive processing, and information is likely to be "stored" in the connections between words, and between words and other sources of information (phonology, semantics) Given this, the multi-component nature of lexical representations leads to a crucial question; does a lexical entry need to be *complete*, possessing both phonology and meaning (and perhaps other things), in order to engage in these processes?

Addressing these questions is important to two major issues in psycholinguistics. The first issue is the nature of lexical representations, and whether there even is such a thing as a single lexical representation, or whether lexical representations are widely

distributed. The fundamental contrast is whether a lexical entry is one complete package, used at multiple levels of the system, or whether different components of a word may play independent roles in different processes. At a more extreme level, it may be that lexical "knowledge" is embedded in a distributed way throughout many processing systems, rather than existing as a piece of stored knowledge. The second issue concerns learning: what processes are necessary for a novel word to become embedded in the processing system? One view is that words need to be learned in full (phonology, semantics, etc) to act as real words in terms of processing. According to this view, simply learning the phonological form of a word should not be enough for it to do things like compete with other words. Alternatively, different processing levels may be able to function partly independently from one another. If this is so, learning the phonological form of a word should be sufficient to enable a newly learned word to compete with other previously known words.

Words play a role at multiple levels of language processing: they influence speech perception (Ganong, 1980; Magnuson, McMurray, Tanenhaus, & Aslin, 2003; McClelland, et al, 2006; Plaut, 2011); they compete with each other during spoken word recognition (Dahan, et al., 2001; Luce & Pisoni, 1998; McClelland & Elman, 1986); they carry syntactic and thematic role information involved in sentence processing (MacDonald, et al 1994; Tanenhaus & Trueswell, 1995), and of course, they refer to objects and concepts (Saussure, 1916; Mayor & Plunkett, 2010). Given this variety of roles a word plays, it has been suggested that simply hearing a string of phonemes a few times is not enough for that string to begin playing these roles: something more complex

may be required for a non-word to begin performing these duties (Storkel, 2001; Gupta, 2003; Gaskell & Dumay, 2003; Leach & Samuel, 2007). However, it is possible that not all components of a word are necessary for a word to play the roles mentioned above. In support of this, existing research suggests that some components of lexical representations, like meaning, are not required for a word to play some of these roles (Abbs, Gupta, and Khetarpal, 2008; Gupta and Cohen, 2002; Gupta and Dell, 1999). For example, Gaskell and Dumay (2003) show that meaning is not necessary for a newly learned word to compete at the phonological level with other real words. However, they also suggest that some sort of sleep-based consolidation is necessary for this competition to arise. In fact, across the few studies examining these issues, we find that novel words can be fully learned (e.g., can play these processing roles) only if they go through some kind of process that will result in sufficient *consolidation* (Gaskell & Dumay) or *engagement* (Leach & Samuel, 2007). The purpose of the present study is to reconsider the interpretation of these findings.

To examine these questions, one must first have a clear idea about what processes words engage in. Since our focus here is on the phonological component of a lexical entry, our focus will be upon the processes of spoken-word recognition: activation and competition via bottom-up flow of information; top-down feedback from words to phonemes (Ganong, 1980; Magnuson, et al., 2003; McClelland, et al., 2006); and lateral flow of inhibition between words (Dahan, et al., 2001; Luce & Cluff, 1998). In the end, by understanding what words do, we can use it to ask our broader questions of 1) whether

all individual components are required for a word to engage in these processes, and 2) more broadly, what it takes for a new word to be learned.

<div align="center">Spoken word recognition</div>

The process of spoken word recognition allows an indirect look at the nature of the lexical entity, as word recognition must operate over links/connections between words as well as between words and lower level representations like phonemes. In many models of spoken word recognition (McClelland & Elman, 1986; Luce & Pisoni, 1998), these connections (the pathways over which processing occurs) are where information about a word is stored. For example, the fact that the word *cat* contains /k/, /æ/ and /t/ is represented by connecting the lexical unit for *cat* to each of these phonemes, such that when the phonemes are activated by the input, the corresponding word becomes active. Similarly, as words become active, feedback connections enable them to affect phoneme perception (e.g., *cat* boosts activation for its constituent phonemes); and lateral inhibitory connections between words allow *cat* to suppress activation for close competitors (*cap, cash*). This provides a powerful way to conceptualize and test learning – since these connections are laid down during learning, and we have unique behavioral markers for many of them, we may be able to use these markers to determine the conditions necessary to establish them. We next present a brief overview of the major findings fielding spoken word recognition, in order to establish measures of these processes, and to discuss how the corresponding connections are laid down during learning.

Parallel activation: It is commonly agreed that spoken word recognition unfolds incrementally over time. Because of this fact, there are points in time in which multiple

words are consistent with the signal heard thus far, creating a form of temporary ambiguity. For example, after hearing just the /kæ/ of cat many words (*catalog*, *can*, *captain*, *cap*, *cat*) are consistent with the input. There is considerable consensus that during this time these potential competitors will be active in *parallel* (Grosjean, 1980; Marslen-Wilson, 1987; Allopenna, Magnuson & Tanenhaus, 1998). As the input unfolds further, this set of candidates is updated such that words that are no longer consistent with the input are eliminated from consideration. Thus, at each point in time the number of words being considered and their degree of activation is changing as more information arrives. Typically by the end of the word the information is sufficient for the listener to successfully recognize the target word. This view of spoken word recognition has been formalized in terms of the Cohort model (Marslen-Wilson, 1987; Marslen-Wilson & Welsh, 1978) according to which at any point in time the only words that the listener considers are those that are consistent with the information presented up to that point. Eventually, only one word is consistent with the incoming information, and is recognized. This point is the *uniqueness point* of the specific word.

Substantial experimental results support this conceptualization (Grosjean, 1980; Marslen-Wilson, 1987; Marslen-Wilson & Zwitserlood, 1989; Allopenna et al, 1998). However this account is not complete. One problem is that listeners are able to recover from temporary misperceptions. For example, listeners might mistakenly parse word boundaries incorrectly (e.g. parse /tulıps/ as tulips instead of *two lips*). Similarly, listeners may mistakenly perceive one phoneme as another (e.g. hear *pone* instead of *bone*). In both cases, according to the Cohort model these misperceptions should be catastrophic,

since there is no way of going back in time and reconsidering any word that was previously excluded as a candidate. However, behavioral results show that listeners can keep this information available and recover if necessary (Gow & Gordon, 1995; McMurray, Tanenhaus, & Aslin, 2009). A second issue is that, in contrast to predictions of the Cohort model, rhymes competitors, which mismatch from the target word at onset, are also active (Connine, Blasko & Titone, 1993); Allopenna et al., 1998). Lastly, Cohort assumes some sort of phonemic input as the basis of word recognition. In this way, subphonemic information (acoustic differences that give rise to the same phonemic percept) may be "lost" at the phoneme level and does not reach the lexical level directly. However research has shown that small, sub-phonemic information is not lost and it can too influence lexical activation (McMurray, Tanenhaus, & Aslin, 2002; Andruski, Blumstein & Burton, 1994; Utman, Blumstein & Burton, 2000).

In contrast, continuous mapping models, like TRACE (McClelland & Elman, 1986) can account for some of these results while retaining some of the basic concepts of Cohort (but see McMurray, et al, 2009). Continuous mapping models are based on the assumption that listeners are flexible in considering a range of possible segmentations and track the optimal one over time. Crucially, this allows word recognition to proceed without requiring listeners to explicitly segment the speech stream (find word onsets/boundaries) prior to accessing the lexicon, and permits rhymes to be considered despite their mismatching onsets.

Overall, this general principle of parallel activation has been adopted by different models of spoken word recognition such as Cohort (Marslen-Wilson, 1987), TRACE

(McClelland & Elman, 1986) and NAM (Luce & Pisoni, 1998). Parallel activation can be cued by both phoneme-level overlap with the target word, but also by sub-phonemic information, like fine-grained changes in VOT.

It is tempting to use parallel activation as a key marker of lexical involvement in processing. That is, we could ask, whether as new words are learned (with or without semantics), other, existing words in the lexicon start getting activated in parallel, as a result of this learning, when these novel words are heard. However, parallel activation on its own does not provide an unambiguous marker. For example, listeners may activate a word (e.g. *disk*) when they hear a similar non-word (e.g. *disp)*, but this may be no different from the case in which we activate a word (e.g. *disk*) when hearing another word (e.g. *dish*). In both cases, activation for *disk* is driven by the temporary ambiguity, the fact that the early portions of the input (/di/) are consistent with both words. That is, listeners may not need to learn anything about *disp* to show such parallel activation. Therefore parallel activation is a result of the temporal ambiguity of the input, but it does not depend entirely on the lexical status of the input, and therefore it may not be a good measure for evaluating the conditions under which novel words are integrated into the lexicon.

Lexical feedback: Research has also considered feed-back connections between words and phonemes. Feedback allows the activation of words (e.g. *beach*) to influence the activation of smaller units, like phonemes (e.g. /b/). As a result, even if the identity of the phoneme alone is not clear from the bottom-up input alone (e.g. the VOT of /b/ is not

representative for a /b/) it can still be sufficiently activated because of the activation it receives from the lexical representation.

The first evidence for this comes from Warren (1970) who presented participants with sentences that contained certain words in which a speech segment was excised and replaced by either cough or a pure tone (e.g. "*The state governors met with their respective legiXlatures convening in the capital city*" where the symbol *X* corresponds to either the sound of cough or a pure tone). None of the participants reported that any phonemes were missing from the sentence, suggesting that the missing information had been restored. Similarly, Ganong (1980) presented participants with ambiguous stimuli, such as *?ift* in which the initial sound would be something between a /g/ and a /k/ and asked them to identify the initial phoneme. Ganong's results showed that listeners were more likely to categorize phonemes in a way that it is consistent with a known word (i.e. /k/ in *?iss* and /g/ in *?ift*). Lastly, Gaskell and Marlsen-Wilson (1998) found that subjects were less able to perceive assimilation in words than nonwords, thus, supporting the hypothesis that lexical feedback influences the perception of phonemes.

These results have undergone significant debate, and it has been suggested that they may be accounted for as a result of memory biases or post-hoc decision processes: listeners may perceive the ambiguity or missing information in the signal but interpret it in a way that is consistent with the lexical input after the fact (Norris, McQueen, & Cutler, 2000). In response, Elman and McClelland (1988) measured phoneme perception indirectly by investigating how a lexically restored phoneme interacts with neighboring phonemes using *compensation for co-articulation* (CfC) (which refers to the finding that

listeners perceive phonemes differently according to the preceding coarticulatory context). Elman and McClelland found that lexically restored phonemes could engage in compensation for co-articulation. As subjects were unlikely to be consciously aware of these subtle relationships, this supports lexical feedback as a "real", not post-hoc effect (see also Magnuson, et al, 2003; Samuel & Pitt, 2003; though see Pitt & McQueen, 1998). Therefore, the behavioral data support the idea that lexical feedback operates over direct connections from the lexical level to the phoneme level.

In addition, these feedback connections may also participate in perceptual learning. Norris, McQueen, and Cutler (2003) exposed participants to a set of words in which one of two fricatives was always ambiguous. However, this ambiguous sound could be interpreted in a way that is consistent with a real word (e.g., a sound that is ambiguous between /s/ and /f/ in ki? should be interpreted as a /s/ because *kiss* is a word and *kiff* is not). After a brief exposure to such words, participants then categorized these ambiguous sounds in a novel set of nonwords, and showed a shift in their category boundary as a result of this learning. That is, feedback between lexical and phonological levels of processing allows listeners to figure out what the appropriate sound is, and retune their perceptual categories. Most importantly Norris et al (2003) showed this perceptual learning paradigm works only when real words guide the perceptual shift during the initial exposure phase.

Feedback thus may exert important effects on both learning and processing and establishing these connections may be a crucial result of word learning. Thus, the presence of feedback from recently learned word-forms may be a useful as a way to

evaluate whether a new word is engaged in lexical processing. Leach and Samuel (2007) used this to examine word learning. They distinguished between *configuration* and *engagement* as two different levels of learning. Configuration refers to learners acquiring the information associated with the word (e.g. phonological form, grammatical use); while engagement refers to the ability of a word to dynamically interact with other words and sub-lexical representations. They claim that "*phonemic restoration ... appears to involve a lexical representation's activation of a subexical phonemic representation … As such, it is a good example of lexical engagement*". Leach and Samuel examined this in a series of experiments in which they used different tasks to measure these two levels of learning following various learning tasks. Specifically, they evaluated lexical configuration with a speech-in-noise task; and evaluated lexical engagement using the lexically-driven perceptual learning task of Norris et al (2003), an indirect measure of feedback.

This study found evidence for feedback-driven perceptual learning on the basis of newly learned words, but only when the novel words were associated with semantic content. No evidence that newly learned words could drive perceptual learning was found when novel words were presented in isolation, and even adding a production task (without semantics) during learning did not yield engagement. This suggested that lexical engagement can only be achieved when certain criteria have been met, like associating a word with a meaning directly (by linking it to a picture), or by using it in a sentence, so that meaning can be extracted. Thus, one might conclude on the basis of this study that

lexical representations must be complete in all aspects in order to be able to dynamically interact with other words and sub-lexical representations.

This conclusion though may be premature as it is based on the absence of an effect, in a paradigm that requires two things to happen – lexical feedback must change the percept in real-time, and then perceptual learning must result from this. However, the broader idea raised by Leach and Samuel, that words dynamically interact with each other (i.e. compete) and this can be used to understand whether a phonological sequence has been lexicalized, offers a powerful way to frame this question. Overall, the logic behind their design seems promising, but different measures of lexical activation may be better in capturing this kind of effects.

Competition and Lateral Inhibition**:** Beyond connections between phonemes and words, words are also directly connected to each other via inhibitory connections. As mentioned above, multiple words are activated during word recognition, and subsequently compete with each other for activation (Luce & Pisoni, 1998; Marslen-Wilson, 1987; McClelland & Elman, 1986; Norris, 1994), such that words with more activation inhibit words with less activation. This may offer an excellent platform for studying novel word learning in the context of spoken word recognition because 1) activation flow within levels of processing may be more robust than activation flow between levels; 2) it may be possible to study competition effects more directly without relying on second-order effects like perceptual learning.

Investigations of *phonological neighborhood* effects offered some of the first evidence for some kind of active competition between words. A word's neighborhood

refers to the group of words that are phonologically similar to it (Goldinger, Luce, & Pisoni, 1989; Luce & Pisoni, 1998). If words actively compete with each other, words with many neighbors should struggle more in order to achieve sufficient activation compared to words with few neighbors. In support of this, Luce & Pisoni (1998) showed that listeners recognize words with large neighborhoods slower than words with small neighborhoods. Interestingly, Vitevitch and Luce found that nonwords do not show the same pattern (Vitevitch and Luce, 1998; 1999). Based on these findings, they suggest that nonwords, in contrast to real words, do not have these direct links to other words and therefore, they cannot receive any inhibition from their neighbors. In other words, nonwords do not have neighbors in the same sense that words do. So, according to Vitevitch and Luce, a crucial distinction between words and nonwrods is that the first are directly connected to each other, forming a kind of lexical network, whereas the later are associated with similar words only indirectly as a result of sharing sublexical information with them. Thus, inhibitory neighborhood effects may be a useful way to examine the conditions under which newly learned words act word-like.

However, the question remains as to the exact mechanism in which this competition between words is implemented. One such mechanism could be *lateral inhibition*. This mechanism proposes that words inhibit each other directly such that more active words inhibit less active ones. Marlsen-Wilson and Warren (1994) used cross-spliced word sequences to ask whether competition between words operates via lateral inhibition. They created stimuli by combining the final portion of one word (e.g. -b from *job*) with the either 1) initial portion of another token of the same word (i.e. jo$_b$- from

job), 2) the initial portion of a different word (e.g. jo- from *jog*, henceforth *jo$_g$b*), or 3) the

initial portion of a nonword (i.e. jo- from *jod*, henceforth *jo$_d$b*). This resulted in stimuli in

which the vowel provided either correct or misleading information about the following

consonant. Critically, the word-splice and non-word-splice conditions allow us to

examine if the additional activation for a competitor word (e.g., after hearing *jo$_g$b*,

listeners will partially activate the competitor, *jog*) inhibits the target word, *job*). The

authors found that lexical decision RTs did not differ across the three splicing conditions,

suggesting that lateral inhibition cannot account for inter-lexical competition.

However, more recent research, using eye-tracking supports lateral inhibition.

Dahan et al (2001) created similar stimuli to those used by Marlsen-Wilson and Warren

(1994), but they used eye-tracking in the visual world paradigm to measure lexical

activation (Tanenhaus, Spivey-Knowlton, Eberhardt & Sedivy, 1995; Allopenna et al,

1998). This study used the same three splicing conditions as Marslen-Wilson & Warren

(1994); a matching splice, a word splice, a nonword splice. Dahan et al presented these

stimuli along with four pictures; the target, one distracter with the same initial phoneme,

and two unrelated distracters[1]. Participants clicked on the picture corresponding to the

target word after hearing the stimulus, and their fixations to each of the four competitors

were monitored as an ongoing measure of lexical activation. Dahan et al found that the

participants' looks to the target were influenced by the splicing condition; participants

[1] For the target word *neck* there was one distracter with the same initial phoneme (e.g. *nurse*) in order to prevent participants from identifying the target picture from the initial consonant alone, before hearing the subcategorical mismatch in the vowel. The other two distracter words were phonologically dissimilar from the target.

looked significantly more to the target (e.g. *neck*) when the target word was spliced with a different token of itself (*ne$_{ck}$ck*) or a nonword (*ne$_p$ck*) compared to the condition in which it was spliced with the initial part of a different word (*ne$_t$ck*). Based on this finding, they argued that when misleading information is provided initially, words consistent with that information are activated, and suppress activation for other competitors. Later on, when information consistent with the target word arrives, it must compete with the previously activated word and it is recognized more slowly as a result of this competition. In contrast, when no misleading information arrives (the matching condition), or when the misleading information does not result in the activation of a different word, this does not preferentially activate a lexical competitor, and thus does not slow recognition of the target to the same degree. Thus, these results support lateral inhibition as a mechanism of lexical competition.

This process of inter-lexical lateral inhibition may provide a useful platform for investigating word learning. Consider Dahan et al's (2001) findings together with those of Vitevitch and Luce (1998; 1999) who argue that nonwords (in contrast to real words) lack direct inhibitory links to words. According to this, nonwords could begin engaging in lexical processes via inhibition if they formed these kinds of links to known words. This would then enable them to compete with real words via the same kind of lateral inhibition described by Dahan et al (2001).

Gaskell and Dumay (2003) tested this by conducting a series of experiments asking whether and under what conditions newly learned words (e.g. *cathedruke*) can compete with other similar known words (e.g. *cathedral*). Participants were first trained

on novel phonological word forms (without any meaning) and then performed either a lexical decision task or a pause detection task. One of the factors manipulated was the time between the training and the testing tasks. Critically, Gaskell and Dumay (2003) found that inhibitory lexical competition effects were shown only after a significant amount of time had elapsed; 5 days to one week after the training. They suggested that while the phonological form could be mastered quite quickly (configuration, in the language of Leach and Samuel, 2007), engagement (in the form of lateral inhibition) may require a significant consolidation period.

Present study

Evidence suggests that spoken word recognition takes advantage of three types of possibly learned connections: bottom-up and feedback connections between phonemes and words, and lateral connections between words. Of these, parallel activation (phonemes→words) does not seem to be helpful in the investigation of what it takes for letter strings to begin acting as words: nonwords could also lead to the activation of similarly sounding words because, as time unfolds, the input can be temporarily ambiguous. Lexical feedback (Ganong, 1980; Elman & McClelland (1988) may also offer a way to evaluate the integration of novel words into the lexicon; if a phoneme is activated by feedback from a newly learned word then this word is sufficiently lexicalized. However, the measures used in this research area, such as Elman and McClelland's (1988) CfC paradigm, or Leach and Samuel's perceptual learning effect might be too subtle and insensitive to different levels of learning. In contrast, competition

may offer a good evaluation of these issues. If a newly learned word can inhibit other known words, this would suggest that it is engaged/integrated into the lexicon. Gaskell and Dumay (2003) used the lexical decision task to evaluate word integration and concluded that it could only be achieved after significant consolidation. Marlsen-Wilson and Warren (1994) used the same task to examine the effect of misleading sub-phonemic information on the speed of spoken word recognition and they also failed to find evidence for inhibition, even among known words. In contrast, Dahan et al (2001) found evidence for inhibition due to temporary sub-phonemic mismatch, using a more sensitive measure of lexical activation (eye-tracking). This could mean that the Gaskell and Dumay's findings could be attributed to that fact that the lexical decision task may be fairly insensitive to lateral inhibition. Thus, it would be useful to re-evaluate Gaskell and Dumay's (2003) results with a more sensitive, visual world task.

The goal of the present study is to use the methodological paradigm of Dahan et al (2001) to investigate how newly learned phonological sequences interact with known words. Specifically, we aim to examine whether just learning the phonological form of a word can lead to immediate signs of lexical competition on that level alone, independently of other lexical aspects (such as orthographic and semantic). The prior reviewed studies suggest that learning the phonological form alone should not be enough in order to observe any kind of lexical competition. However, it is possible that our more sensitive measure might be able to detect such effects. If this is found, it would suggest that lexical entries can compete with each other even when they are not complete, and

more importantly that the kinds of connections necessary for so-called "engagements" (e.g., lateral inhibition, feedback) may be built from the earliest, stages of word-learning.

To assess this, we ran three experiments. The first experiment was a replication of the Dahan et al (2001) study that aimed at establishing that misleading subphonemic information can impede word recognition using our stimuli and our variant of the task. The second experiment examined whether training on the phonological form of a nonword might enable it to compete with real words. To do this, participants were first trained on a set of novel words, and then we examined inter-word inhibition using the same eye-tracking task, but this time comparing inhibition from novel (i.e. trained) words to known words. This led to some rather unexpected results that may have derived from similarity across our item-sets. Thus, the third experiment replicated it with one difference; we minimized the similarity within the nonword set (i.e. the set of stimuli treated as either nonwords or novel words). This found clear evidence for lateral inhibition from newly learned words without meaning or consolidation.

Experiment 1

This experiment replicated the Dahan et al (2001) experiment using a new set of words and in a slightly different eye-tracking task. It was intended to demonstrate that we can observe lexical inhibition using our setup before examining learning in Experiments 2 and 3.

Method

Design

There was a single visual world paradigm eye-tracking task during which participants heard a spoken word while viewing four pictures of potential referents for that word. Participants clicked on the picture corresponding to the auditory stimulus while eye-movements were monitored.

Experimental auditory stimuli were constructed by cross-splicing word stimuli to create three splice conditions; *matching-splice*, *word-splice* and *nonword-splice* (e.g. ne$_{ck}$ck, ne$_{ck}$t, ne$_p$ck). Twenty such triplets were created and within each of them one of the two words was assigned the role of the target (e.g. *neck*). As for the visual stimuli, each trial consisted of a specific set of four pictures. Specifically, for each of the twenty target words three distracter monosyllabic words were selected; one of them began with the same consonant as the target word (e.g. for the target word *mug,* one of the distracters was *milk*) in order to prevent participants from identifying the target picture from the initial consonant alone, before hearing the sub-categorical mismatch in the vowel. The other two distracters were phonologically dissimilar from the target. Thus, there were a total of 20 sets of four pictures each (i.e. eighty pictures in total). Each of the pictures was used as the target in three different trials each of which corresponded to one of the three experimental conditions; *matching-splice*, *word-splice* and *nonword-splice.* This led to the construction of 240 trials in total (20 triplets x 4 pictures x 3 splice conditions). Out of the 240 trials only the 60 experimental trials (i.e. the trials in which the referent corresponded to a target from the twenty experimental triplets) were used in the analysis.

The same 240 were presented twice in two adjacent blocks, thus resulting in 480 trials in total.

<u>Participants</u>

Seventeen undergraduate students at the University of Iowa participated in this experiment and received course credit as compensation. All were native speakers of English. Two of them were excluded from the analyses due to problematic eye-tracking data.

<u>Stimuli</u>

The items consisted of 20 sets of four words (see Table A1 in Appendix). The first word was the target on experimental trials (e.g., *net*). It was matched with a cohort that overlapped on only a single phoneme (e.g., *nurse*), and two unrelated objects. All items were monosyllabic and had a stop consonant as their offset (/b, p, d, t, k, g/).

Auditory stimuli for the target words were constructed by cross-splicing a word with 1) another recording of itself, 2) another word and 3) a nonword (e.g., *net*, *neck*, *nep*). The selection of the words and non-words for cross-splicing was based on two constraints. Specifically, within each triplet the only difference was the place of articulation of the final stop consonant. Second, words within a triplet always matched in final voicing, (though this feature differed between items). For example, one target was constructed by splicing *neck*, *net*, or *nep*, and another one was created from *mug*, *mud*, and *mub* (see Table A2 in Appendix).

In order to construct these stimuli all three words were first recorded by a male native speaker of American English in a sound-proof room, sampling at 44100Hz. We

then extracted the final stop consonant from the target word. This final sound was spliced

onto the initial portion (up to the end of the vowel) of a different token of the target word

to construct the matching-splice condition (ne$_{ck}$ck); the initial portion of a competing

word to construct the word-splice condition (ne$_t$ck), or the nonword to construct the

nonword splice (ne$_p$ck). Splicing was always conducted at the nearest zero crossing to the

beginning of the closure (marked by the loss of F2 and F3 in the vowel), so that the final

consonant included the closure (plus any voicing or aspiration) and the final release.

For each of the 20 targets, each of the three conditions was constructed in this

way (20 stimuli x 3 splicing conditions). The same procedure was followed for the three

filler items in each set as well. The only difference was that each word was spliced with

itself and two nonwords.

Visual stimuli consisted of 80 pictures (20 triplets x 4 pictures). These were

developed using a standard McMurray lab procedure designed to ensure that each picture

represents a prototypical depiction of the target word (McMurray, Samuelson, Lee &

Tomblin, 2010; Apfelbaum, McMurray & Blumstein, 2011). For each item, several

pictures were downloaded from a large commercial clipart database. One picture was

selected by groups of 3–4 viewers as being the most representative, easiest to identify,

and least similar to the others in the complete set. Next, some images were edited to

remove extraneous components or to alter colors. Finally, each picture was approved by a

senior investigator with significant experience using the visual world paradigm.

Procedure

The experiment took approximately half an hour**.** After undergoing informed consent, participants were seated in a sound-proofed room and were calibrated. Next, participants were familiarized with the pictures that were in the experimental sets, by viewing each of the 80 pictures along with an orthographic representation of their name.

After this, they began the experimental trials. Each trial began with a display of four pictures from one of the item-sets in the locations that they would be in throughout the trial along. They were accompanied by a small blue circle in the center of the screen. This allowed the participants to quickly figure out what pictures were present on that trial and where they would be located, so that eye-movements during the auditory stimuli would be more likely to reflect lexical processing, not visual search. After 500 ms this circle turned red, signaling the participant to click on it with the computer mouse. The red circle then disappeared and one of the auditory tokens was played. Participants clicked on a picture, which ended the trial. There was no time-limit on the trials, but subjects typically responded in less than 2 sec (M = 1358.69 ms, SD = 157.23 ms).

<u>Eye-movement recording and analysis</u>

Eye movements were recorded using an SR Research EyeLink II head-mounted eye tracker. At the beginning of the experiment, the eye-tracker was calibrated using the standard 9 point calibration procedure. A drift correction was performed right after the calibration and it also took place once every 12 trials. Fixations were monitored at 250 Hz (every 4 ms). Raw eye-tracking data were automatically parsed into saccades, fixations, and blinks using the default Eyelink parameters. Subsequently, adjacent

saccades and fixations were combined into "looks". Each look event started at the

saccade onset and ended at the fixation offset (McMurray, Tanenhaus, & Aslin, 2002).

Results

The participants' clicking responses were on average 99.28% accurate across all

trials suggesting that all the words were clearly understood by all participants, despite the

splicing manipulation. The results were significantly more robust for the first block (the

first 240 trials)[2] so we didn't include the second block in our analyses.

We focused our analysis of the eye-tracking data on only the one quarter of the

trials in which the target word was the auditory stimulus. On these, we computed the

proportions (across participants) of fixations to the target picture as a function of time, for

each of the three splicing conditions. We discarded trials on which an incorrect picture

was selected (0.72% of trials per participant). Figure 1 presents the proportions of

fixations to the target picture over time for each splicing condition. Much like Dahan et al

(2001) it shows that matching stimuli showed the quickest rise in looks to the target,

followed by nonword-spliced stimuli, with word-spliced stimuli as the slowest.

As in Dahan et al (2001), we found that fixations between conditions started to

diverge about 600 ms after target onset and extended until about 1100 ms. We thus

computed the average proportion of fixations between 600 ms and 1100 ms as our

---

[2] This makes sense given the fact that by the end of the first block participants might have stopped weighting the co-articulatory mismatch because it is not informative within this task.

dependent variable. This was compared across the three splicing conditions (within-subject) using a one-way ANOVA. This found a significant effect of splicing condition ($F_1(2,32)=27.3$, $p<.0001$; $F_2(2,38)=12.7$, $p<.0001$). Planned comparisons showed a significant difference between the matching- and word-splice conditions ($F_1(1,16)=54.0$, $p<.0001$; $F_2(1,19)=19.1$, $p<.0001$), as well as between the word- and nonword-splice conditions ($F_1(1,16)=10.6$, $p=0.005$; $F_2(1,19)=4.74$, $p=0.042$). In contrast to Dahan et al (2001) a significant difference was also found between the matching- and nonword-splice conditions ($F_1(1,16)=17.2$, $p<.0001$; $F_2(1,19)=11.6$, $p=0.003$).
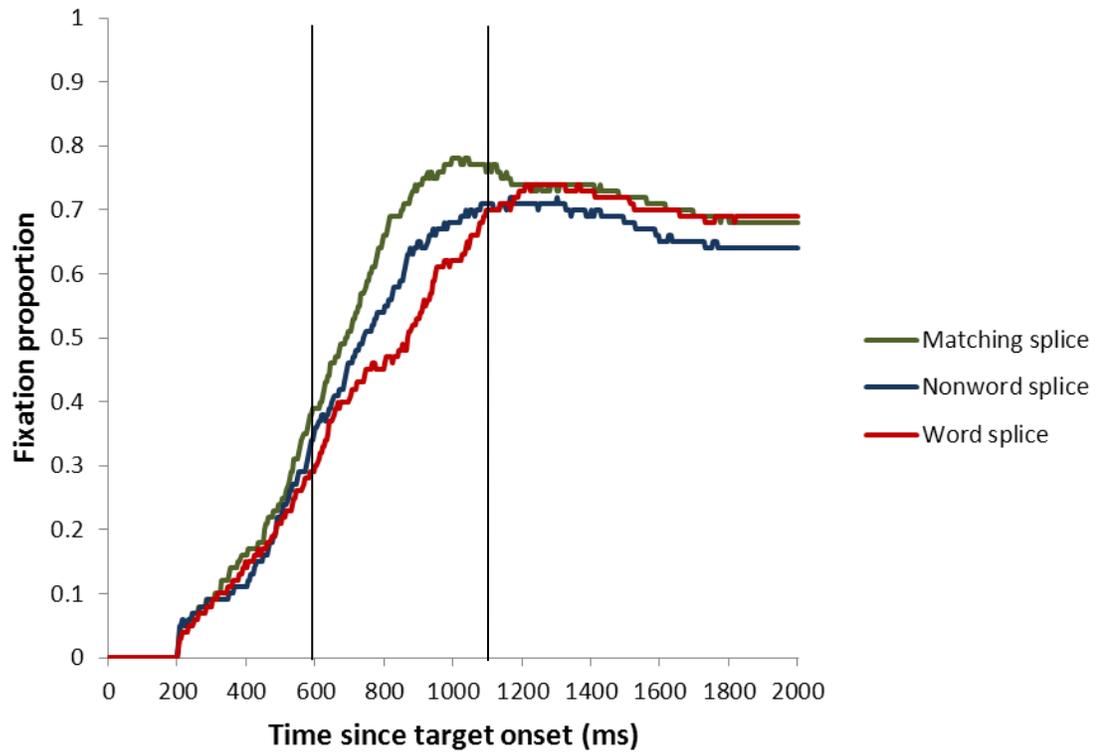
Figure 1. Proportion of fixations to the target for each splicing condition in Experiment 1.

Note: Vertical lines indicate the time window that was included in the analyses

Discussion

Overall, this study replicates the major finding of Dahan et al (2001): sub-phonemic mismatch that contacts another word significantly slows activation over mismatch that does not. This provides continuing evidence for lateral inhibition between words. In contrast to the Dahan et al (2001) we also observed significant difference was also found between the nonword- and matching-splice conditions, suggesting that listeners are also sensitive to fine-grained phonetic differences that do not contact other

words. This finding, however, is in agreement with a number of studies showing a cost to co-articulatory mismatch cost (Gow, 2001; Martin & Bunnel, 1981). Mismatching words are less good exemplars of the target word than matching ones, and as lexical activation is similar to other fine-grained acoustic changes (Andruski et al., 1994; McMurray et al, 2002), the fact that recognition is slower under these conditions was not a great surprise.

Experiment 2 builds on this finding by using these same stimuli in a training paradigm. We trained listeners on the phonological word-form of some of the nonwords from these triplets and then used the paradigm of Experiment 1 to ask if these newly learned words will compete with real words.

Experiment 2

The purpose of this experiment was to examine the effect whether training on novel phonological word forms can allow those word-forms to participate in the lexical inhibition observed by Dahan et al (2001). That is, we wanted to know if a brief exposure to non-words was sufficient to create the lateral connections over which inhibition occurs. Our primary question is whether training on a set of nonwords with *no specific semantic content* allows them to compete with real words at the phonological level. If just experience with the word-form is enough then we will observe a significant difference when targets are spliced with nonwords than when they are spliced from novel words (i.e. trained nonwords). In contrast, if the word-form on its own is not capable of competing with real words then we should observe no difference between these two splice conditions.

Method

Design

Participants first performed a set of two tasks designed to expose them to the nonwords. They were then tested in a visual world paradigm task that was similar to Experiment 1 to assess lexical inhibition after learning.

As in Experiment 1, during the test phase, we used a matching, a word-splice condition as and a nonword-splice condition. This experiment also included a fourth: a *novel-word-splice* condition, corresponding to the words subjects were exposed to during the training phase. Our primary goal was to examine whether any differences would arise between the nonword and novel-word-splice conditions. For each participant, half of the twenty nonwords were assigned to the novel-word condition and half to the nonword condition. The specific nonword set on which participants were trained on was counterbalanced between participants. Specifically, two groups were created (A and B) and within each group half of the participants were trained on 10 nonwords (with the other 10 words being untrained) and the rest of the participants were trained on the remaining ten nonwords. Groups A and B differed in terms of which words were randomly grouped together (see Table 1). Ten participants were assigned to each of the four groups. Participants were first trained on the ten nonwords (training phase) and then performed the eye-tracking task (eye-tracking phase).

Table 1. Sets of nonwords trained in each Group in Experiment 2

Nonwords trained in Group A1	Nonwords trained in Group A2	Nonwords trained in Group B1	Nonwords trained in Group B2
bap	bape	bub	bap
cack	bub	cack	bape
darp	cark	darp	cark
forp	dop	forp	dop
heek	hud	heek	leet
leet	nep	hud	mub
mub	parp	nep	parp
nop	pip	nop	pip
pote	sook	sook	pote
zick	tark	zick	tark

For the training task two tasks were used: 1) a *listen-and-repeat* task, during which participants heard an auditory stimulus and repeated it, and 2) a *stem completion* task, during which participants were given the first part of a word and had to produce the whole word. The whole training session consisted of 11 blocks of each task. The blocks of the two different tasks were presented in an alternating order, always starting with a block of the listen-and-repeat task and ending with a block of the stem completion task. In each block all 10 nonwords (or just the first part of them in the case of the stem

completion task) appeared once. For both tasks, each block consisted of 10 trials, one for each of the 10 trained words, in random order.

The visual world paradigm task was similar to Experiment 1. The biggest difference was that we dropped the second block of trials altogether, resulting in a total of 240 rather than 480 trials. However, the experimental auditory and picture stimuli were identical to that of Experiment 1. At a higher level, however, the half of the items in the nonword-splice condition (those corresponding to the nonwords they were trained on in the training part) were treated as novel-word-splice instead of nonword-splice stimuli. Therefore, each of the 20 experimental words appeared once in the match condition (20 trials), once in the word-splice condition (20 trials), and once in either the novel-word or nonword-splice condition (depending on the subject). This left only 10 trials for each of these two conditions. As before, the other three items in the set were heard an equal number of times for a total of 240 trials (20 item-sets x 4 items/set x 3 splicing conditions). In contrast to Experiment 1, no second block was used, and these were subsequently considered the *novel words* set for that participant.

Participants

Forty-two undergraduate students at the University of Iowa participated in this experiment and received course credit as compensation. All were native speakers of English. Two of them were excluded from the analyses due to problematic eye-tracking data.

Stimuli

The stimuli for the visual world task were identical to that of Experiment 1 (see Table A1). However, a new set of stimuli were constructed for the training tasks. The auditory stimuli for the training phase consisted of the same twenty nonwords of the triplets as in Experiment 1 (Table A2). Specifically, for the listen-and-repeat task whole audio files were used (i.e. without splicing) for each of the nonwords, the specific tokens were different from the ones used in the visual world task. For the stem completion task the auditory stimuli came from separate recordings in which the speaker only spoke the first part of the nonword (e.g. *ne* for the trained nonword, *nep*) so that no coarticulatory information about the following consonant was provided to the participants.

Procedure

The experiment took approximately one hour**.** After undergoing informed consent, participants were seated in a sound-proofed room and were given instructions for the first, training partphase, of the experiment. After completing the training phase participants were familiarized with the pictures for the visual world task, by viewing each of the 80 pictures along with an orthographic representation of their name. Next the eye-tracker was calibrated, participants were given the instructions for the visual world task, and then they proceeded to the experimental phase.

Training phase: The training consisted of eleven epochs of repetition and stem completion. Each epoch consisted of one block of ten *listen-and-repeat* trials and one block of ten *stem completion* trials. During the listen-and-repeat task each word was presented via headphones followed by a cross at the center of the screen. This prompted the participant to repeat the word into a microphone. During the stem completion task,

only the first part of the nonword (e.g., *ne* for the nonword *net*) was presented and participants were instructed to say the whole word. Each nonword was presented twice within each epoch; once in the listen-and-repeat portion and once in the stem completion portion. The order in which the nonwords were presented was randomized within each portion and for each participant separately.

At the beginning of the task each participant was exposed to a practice epoch (consisted of five training items) in order to make sure they had understood the task. During training an experimenter was present in the room in order to make sure the participant performed the task in the correct manner. However, no feedback was given by the computer or the experimenter about the accuracy of the responses.

Experimental Phase*:* The experimental phase of Experiment 2 was the same as the visual world paradigm task used in Experiment 1. Subjects typically responded in less than 2 sec (M = 1481.26 ms, SD = 177.76 ms).

Eye-movement recording and analysis

Eye-tracking data collection and analysis was identical to that of Experiment 1.

Coding of Training Data

Participants' verbal responses in the training phase were transcribed for accuracy by two naïve coders. Mean inter-coder reliability was 95.6%. All the data were included in the analyses irrespectively of any discrepancies between the coders.

Results

In the training phase the participants' average accuracy was 97.15% in the listen-and-repeat trials and 70.08% in the stem completion trials. In the experimental phase they were accurate 98.35% in clicking on the right picture across all trials. Thus, participants learned the words to a fairly high level and had no trouble with these tasks.

To test our primary hypotheses, we examined the experimental phase. Within this, we examined only the experimental trials (i.e. trials in which the auditory stimulus was one of the 20 experimental targets), and only included trials in which participants had clicked on the correct picture (excluding on average 1.86% of the experimental trials per participant).

As in Experiment 1, we computed the proportion of fixations to the target picture over time for each of the four splicing conditions (Figure 2). There is a clear effect of splicing condition with targets spliced with themselves (matching-splice) showing the fastest responding (the quickest increase in looks to the targets). However, the rest of the pattern does not clearly match Experiment 1 or our hypotheses, and subjects appear to be only sensitive to mismatch in general (and not to any differences among the mismatching conditions).
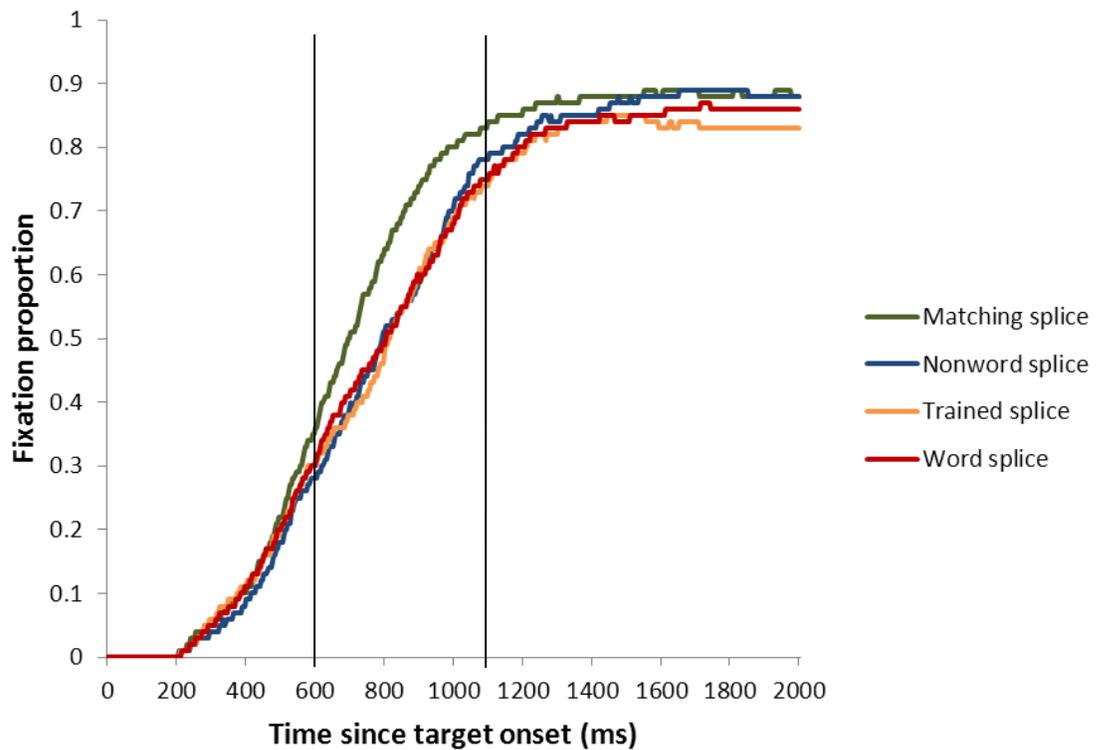
Figure 2. Proportion of fixations to the target for each splicing condition in Experiment 2.

Note: Vertical lines indicate the time window that was included in the analyses

This was confirmed statistically. As in Experiment 1, we computed the average proportion of fixations between 600 ms and 1100 ms as our dependent variable. This was examined in a two-way ANOVA using splicing condition as a within-subjects factor and group (A versus B) as a between subjects factor. This showed a significant effect of splicing condition ($F_1(3,114)=15.3$, $p<.0001$; $F_2(3,57)=7.89$, $p<.0001$). Planned

comparisons revealed that this main effect was due to a significant difference between the matching and word conditions $F_1(1,38)=47.79$, $p<.0001$; $F_2(1,19)=19.46$, $p<.0001$), as well as between the matching and nonword conditions ($F_1(1,38)=35.94$, $p<.0001$; $F_2(1,19)=21.31$, $p<.0001$), but no significant difference between the nonword and novel-word conditions $F_1= 0.76$; $F_2=0.05$, nor between word and nonword conditions $F_1<1$; $F_2<1$. The main effect of group was not significant by subject ($F=1.0$), but it was by item ($F_2(1,19)=10.8$, $p=0.004$). This latter effect was probably due to the fact that different groups of items were used in different training conditions across groups. There was no interaction of group and splicing condition ($F_1<1$; $F2<1$).

Discussion

Overall, the results of Experiment 2 show that people are sensitive to co-articulatory mismatches independently of whether the misleading information results in the activation of a competitor word (both known words and recently learned words). These findings show that training in novel words did not show greater inhibitory effects on the target word in comparison to untrained nonwords.

However, it is important to point out that no difference was found between the nonword-splice and word-splice conditions. These two conditions should have replicated the results of Dahan et al (2001), and more importantly, those of Experiment 1 (since the nonword-splice stimuli of Experiment 2 were a subset of the nonword stimuli of Experiment 1). This raises the possibility of some unforeseen confound in the experimental design. One possible explanation is that fourteen of the 20 triplets shared

their initial consonant with another triplet (e.g. *bait-bake-bape* and *back-bat-bap*). This overlap could lead to partial activation of one when a member of the other was heard. For example if *bape* was trained and *bap* was not, when listening to *ba$_p$ck* (a member of a different triplet), listeners might also partially activate the trained/novel nonword (*bape*) which would then compete with the target word (*back*). In effect, this would blur the line between trained and untrained non-words (and possibly between non-words and words). If this was the case then it should be extremely difficult to be sure about how novel a novel nonword should be considered, given that a very similar nonword had just been subject to extensive training.

This possibility could also account for the lack of significant difference between the nonword and novel word conditions, since the nonword condition might sometimes have indirectly activated similar novel words. In order to test this hypothesis we conducted a second experiment in which we only used nonword and novel word stimuli that did not share their consonantal portion before the vowel.

Experiment 3

The purpose of this experiment was to replicate Experiment 2, but this time after having minimized the amount of indirect activation an untrained nonword might receive from a very similar novel word the participants had just been trained on. Again the main question was whether any significant difference would be found between the nonword and the novel-word-splice conditions.

Method

<u>Design</u>

The design of this experiment was identical to that of Experiment 2. The only difference was in the stimulus set (described in the *Stimuli* section) which were constructed such that there would be no overlap in initial consonants between any of the target words. In addition, we assigned participants to three (A, B and C) instead of two internally counterbalanced groups in order to further minimize any effect of the specific nonword set being treated as trained (i.e. novel word) versus untrained (see Table 2).

Table 2. Sets of nonwords trained in each Group in Experiment 3

Nonwords trained in Group A1	Nonwords trained in Group A2	Nonwords trained in Group B1	Nonwords trained in Group B2	Nonwords trained in Group C1	Nonwords trained in Group C2
bape	brige	bape	darp	chit	bape
cack	darp	brige	forp	forp	brige
chit	forp	cack	grag	grag	cack
grag	jod	chit	leet	heek	darp
heek	nep	heek	nep	jod	leet
leet	parp	jod	parp	mub	nep
mub	sook	mub	rog	sook	parp
rog	stape	stape	shate	stape	rog
shate	tark	tark	sook	weg	shate
weg	zick	zick	weg	zick	tark

Participants

Thirty eight undergraduate students at the University of Iowa participated in this experiment and received course credit as compensation. All were native speakers of English. Two of them were excluded from the analyses due to problematic eye-tracking data.

Stimuli

The construction of the stimuli was identical to that of Experiment 2. The picture stimuli are listed in Table A3 and the auditory ones in Table A4 in the Appendix. In contrast to Experiment 2 none of the triplets shared their initial consonant (except *bait-bake-bape* and *bride-bribe-brige*). All stimuli were recorded by a male native speaker of American English in a sound-proof room, sampling at 44100Hz.

Procedure

The procedure was identical to that of Experiment 2. In the experimental phase subjects typically responded in less than 2 sec (M = 1279.76ms, SD = 150.03 ms).

Eye-movement recording and analysis

Eye-tracking data collection and analysis was identical to that of Experiments 1 and 2.

Coding of Training Data

As in Experiment 2, participants' verbal responses in the training phase were transcribed for accuracy by two naïve coders. Mean inter-coder reliability was 95%. All the data were included in the analyses irrespectively of any discrepancies between the coders.

Results

In the training phase the participants' average accuracy was 96.89% in the listen-and-repeat trials and 71.39% in the stem completion trials. In the experimental phase they were accurate 99.04% in clicking on the right picture. Thus, participants learned the words to a fairly high level and had no trouble with these tasks.

As in Experiment 2, to test our primary hypotheses, we examined the fixation data from the experimental phase, and only the experimental trials (i.e. trials in which the auditory stimulus was one of the 20 experimental targets), in which participants selected the correct picture (excluding on average 2.45% of the experimental trials per participant). As in Experiment 2, we started by computing the proportion of fixations to the target picture over time for each of the four splicing conditions (Figure 3). There is a clear effect of splicing condition with targets spliced with themselves (matching-splice) showing the fastest responding (the quickest increase in looks to the targets). Moreover, in contrast to Experiment 2 further differences were found between the other splice conditions; targets spliced with nonwords showed the fastest responding right after the matching-splice, followed by targets spliced with novel words and words.
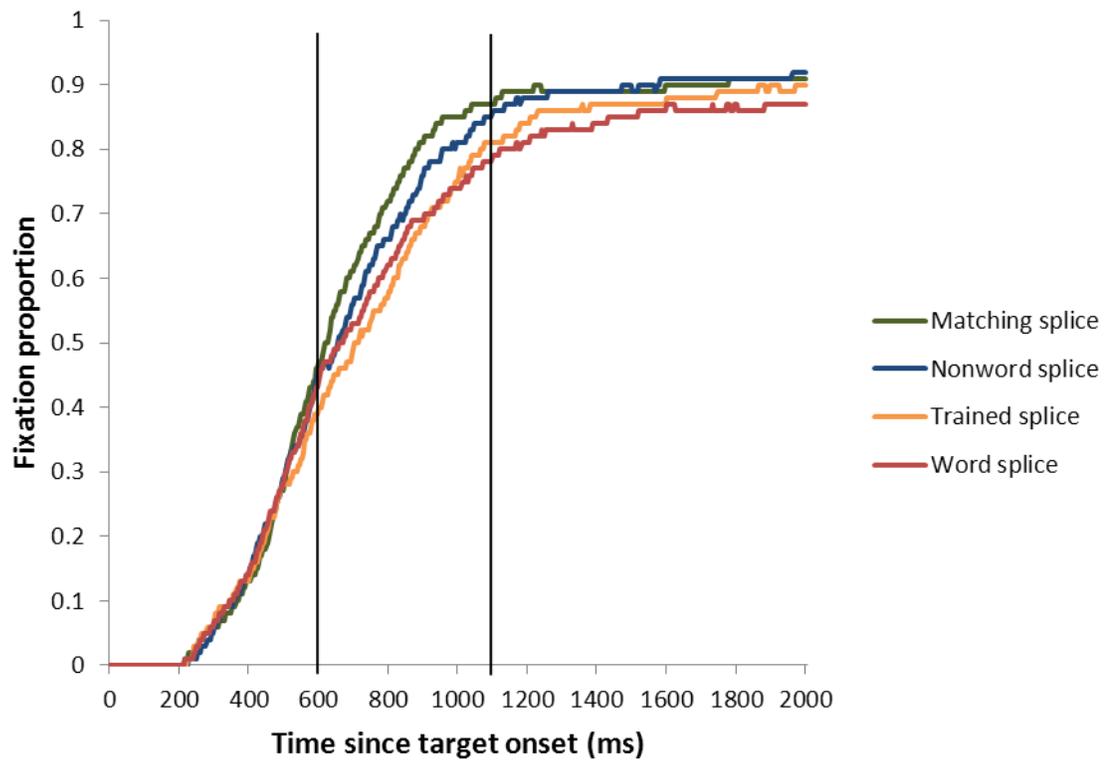
Figure 3. Proportion of fixations to the target for each splicing condition in Experiment 3.

Note: Vertical lines indicate the time window that was included in the analyses

This was confirmed statistically. As in Experiment 2, we computed the average proportion of fixations between 600 ms and 1100 ms as our dependent variable. This was examined in a two-way ANOVA using splicing condition as a within-subjects factor and group (A, B and C) as a between subjects factor. This showed a significant effect of splicing condition ($F_1(3, 99)= 8.63$, $p<.0001$; $F_2(3, 57)=5.50$, $p=0.002$). Moreover, planned comparisons revealed a significant difference between the matching and word

conditions ($F_1(1, 33)=21.09$, P<.001; $F_2(1,19)=12.34$, p=0.002). A significant difference

was found between the matching and nonword conditions in the by subjects ($F_1(1, 33)=$

4.75, p=0.037), but not in the by item analysis ($F_2=1.18$). Most importantly, a significant

difference was found between the nonword and novel word conditions for the by subjects

analysis ($F_1(1, 33)= 4.40$, p=0.044), which was marginally significant for the by items

analysis ($F_2(1,19)=4.36$, p=0.051). In addition, no significant difference was found

between the novel-word-splice and the word-splice conditions ($F_1<1$; $F_2<1$). Finally,

there was no main effect of group ($F_1<1$; $F_2<1$), and no interaction of group and splicing

condition ($F_1=<1$; $F_2=1.27$)

Discussion

At the broadest level, our finding of significant differences between the word and

nonword-splice conditions replicated the results of Dahan et al (2001) as well as those of

Experiment 1 (and unlike Experiment 2). This supports our account of the null effects in

Experiment 2 – training on a set of similar novel-words blurred the line between novel

and non-words, perhaps leading to all of the non-words to begin acting like words.

More importantly, the results of Experiment 3 show not only that people are

sensitive to co-articulatory mismatches, but also that the misleading information may lead

to different effects depending on the kind of stimulus (e.g. real word vs nonword) it

corresponds to. Most importantly, these findings show that training on a novel word form

has an effect on the extent of the co-articulatory mismatch cost. That is, a briefly trained

novel-word can inhibit or slow recognition of a target word, over and above the effect of

the co-articulatory mismatch. Importantly this cannot be attributed to co-articulatory differences in the items themselves (e.g., some words have stronger co-articulation than others) – across subjects the same items appeared in both conditions.

Recall that in Experiment 1, when a listener hears a stimulus like ne$_{ck}$t, which has been cross-spliced with a real word, the competitor word *neck* gets a boost of activation and inhibits the target word *net.* The results show that briefly trained novel-words can achieve the same effect: when a listener hears a stimulus like *ne$_p$ck*, which is the result of splicing a meaningless word-form with a real word, the recognition of the word *neck* will be affected by whether they have been trained on the meaningless word-form *nep* or not (note in our design the stimulus is exactly the same in both cases, across subjects). Specifically, when the listener has been trained on the meaningless word-form the recognition of the real word will be slower. This effect can be explained by means of lateral inhibition; when the misleading information leads to the activation of a novel word-form (*nep*) then this novel word competes with the known word (*neck*), thus slowing down its activation. In contrast, when the misleading information does not result in the activation of a word-form, then the recognition of the real word proceeds with a significantly smaller delay. This would mean that during training a novel phonological word-form representation is formed in the absence of a semantic referent and more importantly, that these inhibitory links can be formed as well by this training, allowing this newly formed representation to compete with other words.

GENERAL DISCUSSION

The present study examined whether learning a phonological word-form, in the absence of both meaning and sleep-based consolidation affects its ability to compete with other words. Leach and Samuel (2007) found that novel phonological word-forms can only develop the feedback connections necessary to support lexically-based perceptual learning if they are learned with a semantic referent. Similarly, Gaskell and Dumay (2003) found that novel phonological forms can start competing with other words only a significant amount of time after they are learned. Based on this finding they suggested that new words form lateral connections to other words only after having been consolidated and that this consolidation takes some time. We tested the necessity of these conditions by using a more sensitive measure of lexical engagement (in our case, in terms of inhibition) based on the visual world paradigm.

In order to do this we used the same design as Dahan et al (2001) to examine whether words inhibit each other during recognition. Dahan et al (2001) found that when mismatching co-articulatory information leads to the activation of another competing real word, recognition of the target word is slowed down. That is, when the mismatching information leads to the activation of another word then this word competes with the word that is being recognized via lateral inhibition. In contrast, when no other word is activated, then no competition takes place and the recognition process is not delayed as much (though it may be delayed by the mismatching perceptual input alone, as we observed in Experiments 1-3). We used the Dahan et al (2001) design but added a

nonword learning task before it to examine whether the learned nonwords would inhibit a target word differently from the novel nonwords (and from the real words). Gaskell and Dumay's (2003) findings suggest that we should not have observed any difference between the novel and the trained nonwords under these conditions, since there was not enough time for consolidation. Similarly, Leach and Samuel (2007) also predict no difference as these words were not trained with any semantics.

It is possible that this is driven by mere exposure to the novel words. This is why we chose to include in the statistical analyses data for all of the novel words independently of how well they were learned by the participants. More likely, however, poorly learned novel words would not yield the same inhibitory effects, suggesting that our inclusion of all the words may have been a fairly conservative choice. However, further investigation should be done. In particular we need to determine if the degree of learning a novel word correlates to the degree in which it interacts with other known words. In other words, if learning a nonword better means that it forms stronger connections to other words then the better it is learned the stronger a competitor it becomes. Therefore, the accuracy scores in the training task would be able predict the amount of delay in recognizing the target word.

The results from Experiment 1 replicated the basic finding of Dahan et al (2001). Specifically, we found that the word-splice stimuli delayed word recognition, whereas the matching-splice stimuli did not, which indicates that there is lateral inhibition between these two words. In addition, we found that nonword-splice stimuli also delayed word recognition, which shows that misleading information disrupts word recognition at some

extent even when no competition takes place at the lexical level. However, this effect is significantly smaller than that of lateral inhibition, as indicated by the difference between the nonword-splice and the word-splice conditions.

The results from Experiment 2 were inconclusive because of the nonword stimulus set; the indirect activation of some novel words, due to the presentation of similar nonwords, blurred the line between the two splice conditions, as well as between the nonword-splice and the word-splice conditions. This was evident from the insignificant differences between these three conditions. This is why we used a different nonword set in Experiment 3.

The results from Experiment 3 replicated those of Experiment 1 as far as the three splice conditions that were shared between experiments are concerned (i.e. the matching-splice, the nonword-splice and the word-splice). Most critically, we found a significant difference between the nonword-splice and the novel-word-splice conditions. Even though, no difference was found between the word-splice and the novel-word-splice conditions, however, the difference between the nonword-splice and novel-word-splice conditions did not persist for as long as the difference between the nonword-splice and the word-splice conditions (as showed in Figure 3) and this is something that might be due to the fact that the new competitors have not formed as strong inhibitory connections with other words. Lastly, a crucial point here is that, in contrast to the Dahan et al experiment, the stimuli were identical between the two critical conditions (i.e. the nonword-splice and the novel-word-splice conditions had exactly the same stimuli). Therefore, any difference between these two experimental conditions can only be

attributed to the experimental manipulation (training) and not to the stimuli themselves, which makes our argument even stronger.

The results from Experiment 3 suggest that learning the phonological form alone enables a novel word to interact with other known words. Therefore, there is no need for a lexical representation to be complete (i.e. entail orthographic, semantic etc information) and it does not need to be consolidated over a long time period for it to form connections to and compete with other words. As stated in the introduction, this is important for two reasons. First, it has important implications for our definition of a lexical representation. These findings suggest that a word could be defined as any phonological sequence we have experience with. Second, it shows that learning a new word is simpler that what has been suggested. No time-consuming consolidation is required for a novel word to start competing with others, nor is meaning required. Rather, words appear to be engaged in the inhibitory network after relatively brief exposure, and it does not seem likely to us that this process needs to be separated from the process of learning the bottom-up connections between phonemes and words (what Leach and Samuel, 2007, term configuration). Rather, the result of such learning may simply be more difficult to detect using coarse grained measures like lexical decision.

Overall, our results show that a word does not have to be complete in all its aspects. A word could just be the conglomeration of and the interaction between different pieces of information, some of which may or may not be present.

REFERENCES

Apfelbaum, K., Blumstein, S.E., & McMurray, B. (2011). Semantic priming is affected by real-time phonological competition: Evidence for continuous cascading systems, *Psychonomic Bulletin and Review*, *18*, 141-149.

Abbs, B., Gupta, P. & Khetarpal, N. (2008). Is overt repetition critical to word learning? The role of overt repetition in word learning with and without semantics. *Applied Psycholinguistics, 29*, 627-667.

Allopenna, P.D, Magnuson, J.S. & Tanenhaus, M.K. (1998). Tracking the time course of spoken word recognition: evidence for continuous mapping models. *Journal of Memory and Language, 38*, 419-439.

Altmann, G.T.M. & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73*, 247-264.

Andruski, J.E., Blumstein, S.E., & Burton, M. W. (1994). The effect of subphonetic differences on lexical access. *Cognition, 52,* 163-187.

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

Dahan, D., Magnuson, J.S., Tanenhaus, M.K., & Hogan, E. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes, 16*(5/6), 507-534.

Elman, J.L., & McClelland, J.L. (1988). Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language, 27*, 143-165.

Elman, J. L. (2009). On the meaning of words and dionosaur bones: Lexical knowledge without a lexicon. *Cognitive Science, 33*, 1-36

Ganong, W.F. (1980). Phonetic Categorization in Auditory Word Recognition. *Journal of Experimental Psychology: Human Perception and Performance, 6*(1), 110-125.

Gaskell, M. G., & Dumay, N. (2003). Lexical competition and the acquisition of novel words. *Cognition, 89*, 105-132.

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

Gow, D.W., & Gordon, P.C. (1995). Human perception and performance. *Journal of Experimental Psychology, 21*(2), 344-59

Grosjean, F. (1980). Spoken word recognition processes and the gating paradigm. *Perception and Psychophysics, 28*, 267-283.

Gupta, P. (2003). Examining the relationship between word learning, nonword repetition, and immediate serial recall in adults. *Quarterly Journal of Experimental Psychology (A), 56,* 1213-1236.

Gupta, P. & Dell, G.S. (1999). The emergence of language from serial order and procedural memory. In B. MacWhinney (Ed.), *The Emergence of Language,* 28th Carnegie Mellon Symposium on Cognition. Hillsdale, NJ, Lawrence Erlbaum.

Gupta, P. & Cohen, N.J. (2002). Theoretical and computational analysis of skill learning, repetition priming, and procedural memory. *Psychological Review, 109*, 401-448.

Gupta, P., & Tisdale, J. (2009). Does phonological short-term memory causally determine vocabulary learning? Toward a computational resolution of the debate. *Journal of memory and language, 61*, 481-502.

Leach, L., & Samuel, A. (2007). Lexical configuration and lexical engagement: When adults learn new words. *Cognitive Psychology, 55*(4), 306-353.

Luce, P.A., & Cluff, M.S. (1998). Delayed commitment in spoken word recognition: Evidence from cross-modal priming. *Perception & Psychophysics, 60*(3), 484-490.

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

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). Lexical nature of syntactic ambiguity resolution. *Psychological Review, 101*, 676-703.

Magnuson, J. S., McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2003). Lexical effects on compensation for coarticulation: The ghost of Christmash past. *Cognitive Science, 27*(2), 285-298.

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

Marslen-Wilson, W.D., & Warren, P. (1994). Levels of perceptual representation and process in lexical access: Words, phonemes, and features. *Psychological Review, 101*, 653-675.

Marslen-Wilson, W.D. & Welsh, A. (1978). Processing interactions and lexical access during word-recognition in continuous speech. *Cognitive Psychology, 10*, 29-63.

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.

Martin, J. G. & Bunnell, H. T. (1981). Perception of anticipatory coarticulation. *Journal of the Acoustical Society of America, 69*, 559-567.

Mayor J. & Plunkett K. (2010). A neuro-computational account of taxonomic responding and fast mapping in early word learning. *Psychological Review, 117*(1), 1-31.

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

McClelland, J.L., Mirman, D., & Holt, L.L. (2006). Are there interactive processes in speech perception? *Trends in Cognitive Sciences, 10*(8), 363–369.

McMurray B., Samelson V.M., Lee S.H., & Tomblin J.B. (2010). Individual differences in online spoken word recognition: Implications for SLI. *Cognitive Psychology, 60*(1), 1-39.

McMurray, B., Tanenhaus, M., & Aslin, R. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition, 86*(2), B33-B42.

McMurray, B., Tanenhaus, M.K., & Aslin, R.N. (2009) Within-category VOT affects recovery from "lexical" garden paths: Evidence against phoneme-level inhibition. *Journal of Memory and Language, 60*(1), 65-91

Norris, D. (1994). Shortlist: A connectionist model of continuous speech recognition. *Cognition, 52*(3), 189-234.

Norris, D., McQueen, J.M., & Cutler, A. (2000). Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences, 23*(3), 299-370.

Norris, D., McQueen, J.M., & Cutler, A. (2003). Perceptual learning in speech. C*ognitive Psychology, 47*(2), 204-238.

Pitt, M.A., & McQueen, J.M. (1998). Is compensation for coarticulation mediated by the lexicon? *Journal of Memory and Language, 39*, 347-370.

Plaut, D.C. (2011). Connectionist perspectives on lexical representation. In G. Gaskell and P. Zwitserlood, (Eds.), L*exical Representation: A Multidisciplinary Approac,* (pp 149-170). New York : De Gruyter Mouton.

Samuel, A.G., & Pitt, M.A. (2003). Lexical activation (and other factors) can mediate compensation for coarticulation. *Journal of Memory and Language, 48*, 416-434.

Storkel, H. L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research, 44*, 1321–1337.

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence Comprehension. In P. D. Eimas & J. L. Miller (Eds.), *Handbook in perception and Cognition. Speech, language and communication.* (Vol. 11, pp. 217-262). New York, NY: Academic Press.

Tanenhaus, M.K., Spivey-Knowlton, M.J., Eberhard, K.M. & Sedivy, J.E. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268*, 1632-1634.

Utman, J. A., Blumstein, S. E., & Burton, M. W. (2000). Effects of subphonetic and syllable structure variation on word recognition. *Perception & Psychophysics, 62*(6), 1297-1311.

Vitevitch, M.S. & Luce, P.A. (1998). When words compete: Levels of processing in spoken word perception. *Psychological Science, 9*, 325-329.

Vitevitch, M.S. & Luce, P.A. (1999). Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory & Language, 40*, 374-408.

Warren, R.M. 1970. Perceptual restoration of missing speech sounds. *Science 167*, 392–393.

APPENDIX

Table A1. Picture stimuli used in Experiments 1 and 2

Target word	Cohort distracter	Dissimilar distracter 1	Dissimilar distracter 2
bait	bed	reed	toad
bat	book	shuck	moat
bud	boat	cord	head
carp	coat	hoop	shack
cat	cook	web	chart
dark	deck	shake	wig
dot	duck	cheek	yacht
fork	feet	goat	wood
heap	hood	shed	fog
hub	harp	rook	shark
knot	note	check	jug
leap	leg	chalk	tub
mug	mitt	lake	yolk
neck	nut	shot	jeep
part	peak	vet	rake
pick	pot	gate	rug
poke	pad	guard	jet
suit	soap	yard	fig
tarp	tape	root	sheep
zit	zap	sword	rock

Table A2. Splicing triplets used in Experiments 1 and 2

Matching-splice	Word-splice	Nonword-splice
bait	bake	bape
bat	back	bap
bud	bug	bub
carp	cart	cark
cat	cap	cack
dark	dart	darp
dot	dock	dop
fork	fort	forp
heap	heat	heek
hub	hug	hud
knot	knock	nop
leap	leak	leet
mug	mud	mub
neck	net	nep
part	park	parp
pick	pit	pip
poke	pope	pote
suit	soup	sook
tarp	tart	tark
zit	zip	zick

Table A3. Picture stimuli used in Experiment 3

Target word	Cohort distracter	Dissimilar distracter 1	Dissimilar distracter 2
bait	boot	head	jug
bride	bread	yacht	vote
cat	cord	beard	blood
chick	check	pig	hook
dark	dog	cloud	ride
fork	fog	god	side
grad	gripe	boat	stork
heap	hood	yard	maid
job	jet	duck	book
leap	lark	wig	peg
mug	milk	truck	spark
neck	nut	goat	wet
part	pad	trout	black
rod	root	feet	bet
shake	shed	keg	dead
state	stick	red	chart
suit	sword	kid	reed
tarp	toad	vet	jeep
web	wood	drop	cook
zit	zap	raid	feed

Table A4. Splicing triplets used in Experiment 3

Matching-splice	Word-splice	Nonword-splice
bait	bake	bape
bride	bribe	brige
cat	cap	cack
chick	chip	chit
dark	dart	darp
fork	fort	forp
grad	grab	grag
heap	heat	heek
job	jog	jod
leap	leak	leet
mug	mud	mub
neck	net	nep
part	park	parp
rod	rob	rog
shake	shape	shate
state	soup	sook
suit	steak	stape
tarp	tart	tark
web	wed	weg
zit	zip	zick