2017

Adult statistical word segmentation across two speakers

Hosung Kang
Western University, hkang72@uwo.ca

Adult statistical word segmentation across two speakers

Hosung Kang

Western University

Supervisor: Dr. Lisa Archibald

Abstract

A finding reliably demonstrated in past research is that statistical learning mechanism facilitates the process of learning language. What remain poorly understood are the effects of multiple speakers in infants and adults learning a statistical artificial language. This study sought to examine the effects of two different speakers in adults because previous literature has suggested that infants lack the ability to segment words when the speech stream consists of two different speakers. Therefore, our experiment sought to understand if 1) adults could successfully segment words across two different speakers and 2) if they can generalize segmentation to a novel voice. Contrary to the infant study, it was found that adults could successfully segment and identify words even when exposed to different speakers. However, adults had difficulty in generalizing to a novel voice when exposed to a single talker. These results support the role of the exemplar theory and raise the possibility that adults are not that experienced language processors as previously expected.

Introduction

Statistical learning is the process of detecting probabilities in the environment to make accurate predictions and to form future expectations (Hasher & Zacks, 1984). By this mechanism of learning, adults and infants can successfully segment words by utilizing transitional probability cues, which is defined as the probability of syllable Y occurring given the syllable X (Romberg & Saffran, 2011; Saffran, Newport, & Aslin 1996). The logic behind statistical learning via detecting transitional probabilities is that when adjacent syllables co-occur frequently, it suggests syllables belonging to the same word. On the other hand, when transitional probability is low, adjacent syllables do not co-occur frequently and suggests a word boundary. Take for example the phrase *pretty baby*. In infant guided speech, syllables "pre" and "ty" co-occur approximately 80% of the time. However, the probability of syllables "ty" and "ba" co-occurring, as in the phrase "pretty baby," is only around 0.03% (Graf Estes & Lew-Williams, 2015; Saffran, 2003; Xie, 2012). Given these variable statistical cues, *pretty* is more likely to be a word and *tyba* to be a word boundary.

To show that adults and infants utilized transitional probability cues to segment words, Saffran et al. (1995; 1996) created an artificial language that consisted of six trysyllabic words. A voice synthesizer was used to remove pauses between words, thereby creating a continuous monotone speech stream where transitional probabilities were the only reliable cues that could be used to segment words. Upon completion of the listening phase, adults performed a two-alternative forced choice test, where participants were given a choice between a non-word and a word from the language. Results suggested that they were able to correctly identify words from the language, above chance. Similarly, infants were placed in a head-turn preference procedure, which measures how long an infant fixates on a visual cue when an auditory stimulus is played over the background. This difference in time of visual fixation determined the preference of sounds, and indeed infants showed preference to novel words rather than words from the language. Thus, it was suggested that infants were able to utilize transitional probability cues to correctly segment and identify words from the language.

Statistical language learning was often criticized that it does not hold up to the demands found in natural environment because artificial language paradigm was tested using a voice synthesizer (Johnson & Tyler, 2010). To prove otherwise, Graf Estes and Lew-Williams (2015) employed the same head-turn preference procedure used in Saffran’s study, but infants were exposed to a continuous speech stream consisting of natural multiple voices, rather than a monotone synthesized voice. Indeed, natural spoken language has variations in speaker identity, tone, pitch, affect, and rate, which a synthesized voice does not provide (Singh, White, & Morgan, 2008). Regardless, results suggested that infants were able to differentiate between words when the speech stream consisted of eight different speakers. However, when the speech stream consisted of two different speakers, infants were not able to differentiate between words. This evidence was surprising, given the fact that infants in Saffran’s study (1996) were able to show a difference in word preference in as little time as two minutes, but infants in the aforementioned study could not differentiate at all, even though the exposure time was in excess of nine minutes. Given these contradictory results, the purpose of this study was to determine if such effect is found in adults learning the same artificial language paradigm.

Rationale

According to the exemplar-based model of speech perception, speaker-specific information, such as the characteristics of the voice, is not discarded but kept in the same memory where words are stored (Goldinger, 1998; Goldinger Pisoni, & Logan, 1991). Because this information is kept within one's lexicon, retrieving a word from memory is easier when the voice cue is similar to the voice that was heard when learning took place. The voice cue in itself activates the memory trace of words spoken by the specific speaker (Nygaard & Pisoni, 1998; Palmeri, Goldinger, & Pisoni, 1993). For example, if a student were to learn the word *hypothesis* from a professor, the professor's voice would be stored in that student's memory of the word *hypothesis*. Consequently, memory retrieval of that word becomes easier when the student hears the same professor's voice. Indeed, previous experiments have shown, in both infants and adults, that word recognition performance degrades when voice changes from when learning took place to the time of word recall (Goldinger, Pisoni, & Logan, 1991; Houston & Jusczyk, 2003). Conversely, word recognition performance remains high when the same voice persists over learning and test phase (Mullennix, Pisoni, & Martin, 1989). Thus, if speaker-specific information plays a prominent role in statistical language learning, the group condition that has the same speaker for listening and test phase should have the highest word identification performance.

While there is some evidence supporting the effect of speaker-specific information and word recognition, others argue that it is implausible to store every speech ever heard as its own representation due to insufficient memory storage (Johnson, 2005). Furthermore, when one hears the word *baby* from two different speakers, each speech is normalized such that perception of the word remains the same despite its large sound variance (Syrdal & Gopal, 1986). Thus, by this view, multiple voices are filtered during the learning process and word identification is spared due to speech normalization. Also, statistical learning has been observed across domains such as visual, audiovisual, tones, and across species (Fiser & Aslin, 2001; Kirkham, Slemmer, & Johnson, 2002; Newport, Hauser, Spaepen, & Aslin, 2004; Saffran, Johnson, Aslin, & Newport, 1999). Given the fact that even primates and non-primates, such as rats, are able to compute statistical regularities in speech input, this suggests that statistical learning is a generalized learning process where probability cues are the determining factor as to whether learning takes place or not (Hauser, Newport, & Aslin, 2001; Toro & Trobalón, 2005). While there may be some effect of voice cue facilitating word recognition, statistical language learning is solely guided by detecting transitional probabilities, and as such, inhibition of learning due to variations in voice would be insignificant and not be observed in adults.

Hypothesis

Adults utilize transitional probability cues to correctly segment and identify words even when surface variation is introduced.

Prediction

If adult participants were to listen to a continuous stream of artificial language for eight minutes with alternating female and male voice guided only by transitional probability cues, they will successfully identify words from the language regardless of surface form variations.

Method

Participants

Participants in the present study consisted of 80 adults (M_{age} = 18.94 years, SD_{age}

= 15.06, N_{male} = 30, N_{female} = 50). All participants were recruited from the undergraduate psychology pool at Western University and received course credit for study completion. All subjects reported English as their first language and two participants had hearing difficulties in one ear. Two Participants did not identify English as their first language and were excluded from the study.

Procedure

Upon arrival, participants received a letter of information, signed consent form, and completed a short questionnaire to obtain the following demographic information: age, gender, first language, number of years speaking English, and vision or hearing difficulties. All participants were tested in a quiet room where the task was administered via a laptop. Participants were quasi-randomly assigned to one of the four conditions in a sequential order with no participant factor determining group assignment. There were equal numbers of participants between groups as data collection progressed. After participants filled out their questionnaire form, they immediately went into the listening phase to be exposed to the artificial language. After completing the listening phase, participants completed the test phase. Upon completion of the study, participants received a debriefing form detailing the experimental manipulation.

Artificial language stimuli

The language consisted of four consonants (d, b, p, t) and three vowels (i, a, u) where combinations of consonant-vowel pairs made up 12 syllables. These syllables were then combined to make six trisyllabic words: *dutaba, bupada, tutibu, patubi, pidadi*. Some syllables occur in more than one word and thus, syllables within a word had transitional probability ranging from 0.33 to 1.0, whereas a word boundary had transitional probability ranging from 0.1 to 0.2. For example, the syllable *bu* occurred in three words, whereas *ti* occurred in one word.

Recording the artificial language

There were two audio recordings of the language that contained the exact same sequence of words, but differed in number of speakers. First recording was created using only one female voice. Second recording was created using the same female voice and an additional male voice that alternated every one-minute. All transition of speakers occurred between words. Both speakers identified English as their native language. Audio recordings of the language were constructed from independent male and female recordings of three-syllable sequence of every articulation within the language. Middle syllable was then excised and concatenated to the language and this manner was carried out to complete the speech stream. For example, to create the language in order of sequence of *tu ti bu du ta ba*, the sequence *tu ti bu* was recorded and the middle syllable *ti* was excised and implemented into the language. If the target syllable was *bu*, the sequence *ti bu du* was then recorded and the syllable *bu* was excised and implemented into the language. There were no pauses between syllables or words, and as such, created a continuous natural flow of the artificial language. In total, there were 140 tokens of each word where the same word never occurred twice in a row. This created the artificial language stimuli that lasted 8-minutes.

Listening phase procedure

Participants were quasi-randomly assigned to one of two listening conditions: female-talker or alternating-talker stimuli. Listening phase was administered via a laptop using E-prime 2.08 software. Before each listening phase began, an instruction was displayed on the screen that told the participants that they would hear a nonsense language and that their goal was to figure out where words began and ended. Throughout the listening phase, variations of lower case and upper case letters were randomly displayed on the screen every two seconds. There were no instructions regarding these letters displaying on the screen. The purpose of these letters was to stop participants from dozing off or not paying attention to the task. Once the listening phase was over, participants immediately went into the test phase.

Test phase stimuli

Six non-words were created from the same syllables that were used to create the language: *pubati, tapudi, dupitu, tipabu, bidata, batipi*. Two separate male-voice and female-voice test stimuli were created with the same male and female voice used to create the artificial language. One of these two test conditions was administered after each listening phase, thus creating four separate groups within the study. For example, a participant was either assigned to a female-talker condition or an alternating-talker condition and was tested with a male-voice stimuli or female-voice stimuli. Since these non-words were novel, transitional probabilities of syllables within the word were zero.

Test phase procedure

Immediately after the *listening phase*, participants were tested on their knowledge of the language. The test was a two-alternative forced choice test and was administered auditorily, where each trial contained a non-word paired with a word from the language. These two choices of words were separated by a silence of 500ms. Upon hearing the two choices, participants selected the *A* or *L* key on the keyboard, respectively, to indicate the word that most sounded like something from the language. There were 36 trials, as each word from the language was paired with a non-word.

Analytic procedures

To determine if speaker specific identity had an effect in the level of learning, it was important to first establish that learning of the language took place for each group condition. A single-sample *t* test (two-tailed) was used to determine if participants within the group had identified words from the language above chance. Independent *t* tests were also used to compare between separate groups to determine the effect of speaker advantages and the effect of speaker variability. Scores from the two-alternative forced choice test were the dependent variables.

In addition, each participant was labelled as learners or non-learners. As a group performance, this value may be numerically greater than the threshold of 50%, but not every individual have a score greater than 18 (chance performance). Therefore, this analysis provided a different perspective as to determine how many people within each groups were identified as "learners" that contributed to the above chance performance as a collective group. To be identified as a learner, individual score needed to have their z score greater than 1.645. Z-score was calculated using this formula where x was individual score:

$$z = \frac{\frac{x}{36} \times \frac{18}{36}}{\sqrt{\frac{\frac{1}{2} \times \frac{1}{2}}{36}}}$$

After learners and non-learners were identified within each group, chi square analysis was used to compare between groups in terms of number of learners.

Results

Table 1 presents descriptive statistics for each group condition. As expected, each group performed well above the chance performance. We used an alpha level of .05 for all statistical tests. Although, female-talker/male-test had a score above chance performance, it had the lowest score overall. An independent *t* test comparing female-talker/female-test and female-talker/male-test was significant, $t(36) = 2.131$, $p = 0.04$. However, one-way ANOVA comparing between alternating-talker/female-test, alternating-talker/male-test, and female-talker/female-test was not significant, $F(2, 56) = 0.61$, $p = 0.55$. Therefore, it was suggested that there were speaker-specific advantages in word-identification and that variability in speakers does not necessarily facilitate learning.

Table 1. Mean group word-identification score and significance tests comparing chance performance

Group condition	Mean word-identification score	One-sample *t* test
Alternating-talker/female-test	23.45 (SD = 4.3)	$t(19) = 5.62$. $p = 0.000$
Alternating-talker/male-test	22.85 (SD = 3.1)	$t(19) = 7.08$, $p = 0.000$
Female-talker/female-test	22.16 (SD = 3.4)	$t(18) = 5.43$, $p = 0.000$.
Female-talker/male-test	20.00 (SD = 2.9)	$t(18) = 3.02$, $p = 0.007$

The number of learners and non-learners in each group condition is presented in Table 2. When comparing all four groups, there was a significant chi square value, $\chi^2(3) = 9.261$, $p = 0.026$, meaning there were different distributions of learners and non-learners throughout different groups. The distribution of learners and non-learners did not differ between alternating-talker/female-test, alternating-talker/male-test, and female-talker/female test, $\chi^2(2) = 2.275$, $p = 0.32$. Thus, participants in the female-voice/male-test had more non-learners and fewer learners compared to other groups.

Table 2. Number of learners and non-learners in each group

	Group Condition				
	AT/FT	AT/MT	FT/FT	FT/MT	Total
Learners	35.14 (13)	32.43 (12)	21.62 (8)	10.81 (4)	100.00
Non-learners	17.07 (7)	19.51 (8)	26.83 (11)	36.59 (15)	100.00
Total	25.64 (20)	25.64 (20)	24.36 (19)	24.36 (19)	(78)
Chi Square	DF = 3	Value = 9.26	Prob = 0.026*		

Note: Numbers in parentheses are number of participants within each group. Groups are labelled as follows: AT/FT = alternating-talker/female test; AT/MT = alternating-talker/male-test; FT/FT = female-talker/female test; FV/MT = female-voice/male-test.

Discussion

The purpose of this study was to determine the effect of two different speakers in adult statistical language learning paradigm. We hypothesized that statistical language learning is a general learning process such that variability in speakers would not affect the process of word segmentation and word identification. Indeed, there was no evidence that surface variations inhibited adults learning the artificial language. However, results suggested that adults had difficulty in generalizing to a novel voice stimulus. All groups learned the words from the language above chance performance suggesting participants successfully segmented words within the speech stream by utilizing transitional probability cues, but FT/MT group performed significantly worse than the other groups. Indeed, there were significantly more non-learners as opposed to learners within the group.

Previous literature have shown than speaker specific information may be encoded within the same memory trace of learned words, and subsequently the voice cue in itself can facilitate in word recall (Goldinger, 1996; Nygaard & Pisoni, 1998; Palmeri, Goldinger, & Pisoni, 1993). Therefore in this current study, it was suggested that female-talker information have been encoded within the learned words, thereby giving an advantage in word identification when hearing the same voice. On the other hand, participants in the FT/MT group showed no benefits in encoding speaker specific information, as the novel male voice would not have activated the memory trace of learned words. Furthermore, since both alternating-talker conditions had comparable scores to FT/FT group and also significantly better score than FT/MT, it was suggested that speaker variability had no effect in word identification.

Previous literature that used the artificial language paradigm has shown similar results in that familiarity of speakers facilitates and extremely novel voice inhibits word recognition. However, this familiarity effect was marginal and only mildly supported for the role of speaker specific advantages (Finley, 2013). Furthermore, Voulomanos et al. (2012) employed the same statistical artificial language paradigm where participants were tested on their ability to identify words from the language using the two-alternative forced choice test method (similar to this study). It was suggested that learners were able to identify words from the language after a voice change. In other words, adults could generalize their segmented lexical units to a novel voice. Therefore, results from our study were inconsistent to previous literature in that, adults did not show the same level of generalization. Although adults could successfully generalize to a novel voice in the previous study, word recognition was severely reduced when the same test voice was mildly distorted.

From this, one possible reason for the observed results in our study was that the male voice was largely different, comparable to the level of distortion (see Vouloumanos, Brosseau-Liard, Balaban, & Hager, 2012), such that word-identification was interrupted. However, this was not likely because the same male-test stimuli was administered after alternating-talker phase and still performed significantly better. Therefore, we argue that there were no issues with the male-test stimuli that attributed to the level of word identification. Instead, the differences in performance must have been from the change in voice and lack of ability to generalize.

The present findings can be best understood within the context of Johnson's exemplar theory (see Johnson, 2005). According to this theory, when learners experience a novel exemplar of a syllable, it activates prior exemplars based on similarity. As adults go through the segmentation phase, exemplars of the speaker and segmented words are stored within their memory. Subsequently, words during the test phase can match in familiarity of exemplars, thereby facilitating in word identification. Conversely, when the novel voice exemplar does not match pre-existing exemplars, there is less activation of the memory trace resulting in lower word identification performance. Similarly, exemplars within the alternating condition emphasize the commonality of phonemic information and deemphasize the variable characteristics of the voice. Therefore during testing, characteristics of the voice do not play a prominent role as compared to phonemic properties of the words from the language in word identification.

One potential confound to this study was that in the alternating-talker condition, voice alternated between words rather than within words, thereby creating inadvertent cues as to let the participant know when the words began. Although this effect may be minimal given the fact that in total there were eight alternations, it was still unlikely as this was the reason why these groups performed significantly better than the female-talker/male-test group. Rather, it was most likely due to the fact that adults had difficulty in generalizing. However, future studies could account for this confounding error by alternating speakers within words as opposed to between words thereby not creating a supplementary cue.

Although this experiment did not test the effect of multiple speakers, future studies could build on this existing design to test for the problem of generalization in adults. From this study, the notion that adults have a harder time in generalizing to novel test stimuli was supported. However, it is not yet known how having multiple speakers could effect in the generalizing process. Therefore, future studies could implement multiple speakers to the artificial language paradigm and employ novel test items. Furthermore, in Graf Estes and Lew-William's study (2015), the number of alternating speakers was eight in the first two experiments and two speakers in third and fourth experiment, alternating very frequently (10-15 sec) compared to this study (1 min). Results suggested that infants could learn the words from the language with difficulty, as shown by head-turn preference procedure, when the language consisted of eight different, but could not segment words when the language consisted of two different speakers.

Similarly, future studies could create an artificial language consisting of varying number of speakers and alternating frequently, as opposed to every one minute, to determine the effect of multiple speakers in word segmentation.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*, 321–324.

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities pitting abstract against stimulus-specific representations. *Psychological Science, 17,* 905-912.

Echols, C. H., & Newport, E. L. (1992). The role of stress and position in determining first words. *Language Acquisition, 2,* 189-220.

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, *32*(05), 429–448.

Finley, S. (2013). Generalization to unfamiliar talkers in artificial language learning. *Psychonomic Bulletin & Review*, *20*(4), 780–789.

Fiser, J., & Aslin, R. N. (2001). Unsupervised Statistical Learning of Higher-Order Spatial Structures from Visual Scenes. *Psychological Science*, *12*(6), 499–504.

Frank, M. C., Tenenbaum, J. B., & Gibson, E. (2013). Learning and Long-Term Retention of Large-Scale Artificial Languages. *PLoS ONE*, *8*(1), e52500.

Goldfield, B. A., & Reznick, J. S. (1990). Early lexical acquisition: Rate, content, and the vocabulary spurt. *Journal of Child Language*, *17*(1), 171–183.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*, 251–279

Goldinger, S. D., Pisoni, D. B., & Logan, J. S. (1991). On the nature of talker variability effects on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 152–162.

Graf Estes, K. (2012). Infants generalize representations of statistically segmented words. *Frontiers in Psychology, 3*, 447.

Graf Estes, K. & Lew-Williams, C. (2015). Listening Through Voices: Infant Statistical Word Segmentation Across Multiple Speakers. *Developmental Psychology, 51*(11), 1517-1528.

Hasher, L., & Zacks, R.T. (1984). Automatic processing of fundamental information. American Psychologist, 39, 1372-1388.

Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: statistical learning in cotton-top tamarins. *Cognition, 78*(3), B53–64.

Houston, D. M., & Jusczyk, P. W. (2003). Infants' long-term memory for the sound patterns of words and voices. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 1143–1154.

Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, *13*(4), 341–348.

Johnson, K. (2005). Speaker Normalization in Speech Perception, in The Handbook of Speech Perception (eds D. B. Pisoni and R. E. Remez), *Blackwell Publishing Ltd, Oxford, UK.*

Johnson, K. (2005). Decisions and Mechanisms in Exemplar-based Phonology. *UC Berkeley: Department of Linguistics.*

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: evidence for a domain general learning mechanism. *Cognition, 83*(2), 35–42.

Lew-Williams, C., Pelucchi, B., & Saffran, J. R. (2011). Isolated words enhance statistical language learning in infancy. *Developmental Science, 14*, 1323–1329.

Mcdonald, S., & Ramscar, M. (2001). Testing the distributional hypothesis: The influence of context on judgements of semantic similarity. *In Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 611–616).

Mullennix, J. W., Pisoni, D. B., & Martin, C. S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America, 85*, 365–378.

Newport, E. L., Hauser, M. D., Spaepen, G., & Aslin, R. N. (2004). Learning at a distance II. Statistical learning of non-adjacent dependencies in a non-human primate. *Cognitive Psychology*, *49*(2), 85–117.

Nygaard, L. C., & Pisoni, D. B. (1998). Talker-specific learning in speech perception. *Perception & Psychophysics*, *60*(3), 355–376.

Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 309–328.

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews. Cognitive Science*, *1*(6), 906–914.

Saffran, J. R. (2003). Statistical language learning: Mechanisms and constraints. *Current Directions in Psychological Science*, *12*, 110–114.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926-1928.

Saffran, J. R., Newport, E. L., Aslin, R. N. (1995). Word Segmentation: The role of

Distributional Cues. *Journal of Memory and Language, 35*, 606-621.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52.

Singh, L., White, K. S., & Morgan, J. L. (2008). Building a word-form lexicon in the face of variable input: Influences of pitch and amplitude on early spoken word recognition. Language Learning and Development, 4, 157–178.

Syrdal, A. K., & Gopal, H. S. (1986). A perceptual model of vowel recognition based on the auditory representation of American English vowels. *The Journal of the Acoustical Society of America*, *79*(4), 1086–1100.

Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin, 139*, 792-814.

Thompson, S. P. & Newport, E. L. (2007). Statistical Learning of Syntax: The Role of Transitional Probability. *Language Learning And Development*, *3*(1), 1–42

Toro, J. M., & Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception & Psychophysics*, *67*(5), 867–875.

Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134*, 552-564.

Vouloumanos, A., Brosseau-Liard, P. E., Balaban, E., & Hager, A. D. (2012). Are the products of statistical learning abstract or stimulus-specific? *Language Sciences*, *3*, 70.

Xie, Y. (2012). Transitional Probability and Word Segmentation. *International Journal of English Linguistics, 2*(6).