

# EFFECTS OF LEXICAL FREQUENCY AND LEXICAL CATEGORY ON THE DURATION OF VIETNAMESE SYLLABLES

Marc Brunelle, Daryl Chow and Nguyễn Thụy Nhã Uyên

University of Ottawa

[marc.brunelle@uottawa.ca](mailto:marc.brunelle@uottawa.ca), [dchow055@uottawa.ca](mailto:dchow055@uottawa.ca), [tnguy221@uottawa.ca](mailto:tnguy221@uottawa.ca)

## ABSTRACT

Our study looks at the effect of lexical frequency, lexical categories and phrase boundaries on syllable duration in Vietnamese. We use durational data to shed light on the status of some ambiguous lexical categories such as kinship terms and positional nouns, and to gather additional evidence on the behaviour of some verbs that have grammaticalized homophones. Our results show that high frequency words tend to be shorter, that function words are independently shorter than lexical words, and that Vietnamese has pre-boundary lengthening. They also suggest that, in terms of duration, positional nouns pattern with lexical words, and that pronouns derived from kinship terms and grammaticalized verbs are not durationally distinct from their non-grammaticalized counterparts.

**Keywords:** Southern Vietnamese, corpus phonetics, function words and lexical words, duration, lexical frequency, pre-boundary lengthening.

## 1. INTRODUCTION

Function words and lexical words have been noted to display different phonological and phonetic properties, e.g. [14, 18]. For example, in English, monosyllabic function words are normally unstressed and cliticize to a neighbouring host, while lexical words must be stressed. Function words have also been noted to be subject to phonetic reduction, e.g. [2, 18]. Predictability also has a strong effect on the reduction of both function and content words [6, 12]. In addition, high frequency words have been shown to be more susceptible to phonological reduction and change [3, 8, 9, 12, 17]. However, phonetic differences between function and lexical words are maintained even after controlling for lexical frequency [2].

The question of possible differences between function and lexical words has been raised in a number of studies on phrasal stress and prominence in Vietnamese (VN), but none of these have approached the question from an experimental perspective [6, 7, 10, 11, 15, 16, 19]. Although there is no word stress in VN, which makes it difficult to clearly identify clitics, the general consensus seems

to be that function words are realized with a shorter duration and that their tones have reduced contours [4, 11]. However, it is still unclear if this is merely due to lexical frequency or if this is an inherent property of function words.

A related issue in VN linguistics is the status of some lexical categories [11]. For instance, positional nouns (or relator nouns) like *dưới* ‘under’, *trên* ‘on’, and *trong* ‘in’ can be analyzed either as nouns or function words denoting direction and complementing a verb: *xuống dưới* ‘to go down under’, *ở trên* ‘to be located on/over’, and *vào trong* ‘to enter inside’ [7, 11, 15, 16]. Can phonetic properties be used to determine if these categories are lexical or functional?

More interestingly, some VN words can be used either as lexical or function words. Kinship nouns, for instance, can also be used as pronouns, in which case they seem to undergo reduction [4]. For example, the words *cháu* ‘grandchild’ and *bà* ‘grandmother’ can be used as 1<sup>st</sup> or 2<sup>nd</sup> person pronouns as in the sentence *cháu đưa bà*, which can mean either ‘I bring it to you’ or ‘you bring it to me’ depending on who is talking. Another case is verbs which have grammaticalized variants with a prepositional use, such as *ở* ‘to reside/at’, *với* ‘reach/with’, *cho* ‘give/to’, and *để* ‘put/in order to’ [7, 11, 13, 16]. While some authors adopt the view that these are cases of unique words in different syntactic positions, others, while recognizing that grammaticalization is involved, analyze them as homophones [4, 6, 11, 13].

In this study, we tease apart some of the factors that play a role in durational prominence in VN. We show that durational differences can be used as a diagnostic to address long-standing issues about the nature of some lexical categories in VN. Our aims are:

- 1) to identify potential durational differences between VN lexical and function words,
- 2) to determine to what extent these differences are due to lexical frequency effects,
- 3) to confirm the existence of pre-boundary lengthening [1, 5, 20] in VN, and
- 4) to look at word classes ambiguous with respect to lexical category to determine if their duration correlates with their function. More specifically, we look at positional nouns, and at the kinship terms and verbs that have grammaticalized homophones.

## 2. METHODOLOGY

### 2.1. Corpus, annotation and measurements

A corpus of spontaneous southern VN speech composed of 64,639 syllables was collected (85.8% of words are monosyllabic and 13.6% are disyllabic). It consists of two television interviews (3 speakers) and four conversations with pairs of native speakers (8 speakers). Five speakers are female and six are male, and they were all born between 1949 and 1992. All speakers speak southern dialects, and the speakers selected for conversations were all born and raised in Ho Chi Minh City or the Mekong Delta from southern VN parents. The interviews were downloaded from YouTube and were selected for their good sound quality and the limited amount of overlap between speakers. The conversations were recorded in stereo on a Marantz PMD-671 with two Shure SMD10A head-mounted microphones (one channel per speaker).

The corpus was transcribed and annotated for parts of speech by native speakers under the supervision of the first author. The parts of speech retained for the annotation are positional nouns (P), pronominal kinship terms (K), grammaticalized verbs (G), and all remaining lexical words (L) and function words (F). Because of the low number of polysyllabic function words that could have been used in statistical comparisons, only results from monosyllabic words will be presented here (for this reason, *syllable* and *word* will from now on be used interchangeably). Unintegrated loanwords (414 syllables) were also excluded from the analysis.

The duration of each syllable was automatically extracted using Praat scripts. Twenty-three syllables with durations above 1 second were excluded as outliers.

In the absence of a southern VN frequency database, lexical frequencies were calculated from the corpus and from a smaller corpus composed of two comedy skits totalling an additional 12,017 syllables. The effect of homophony was partly controlled by distinguishing homophones that belong to different lexical categories. Word and syllable frequencies were both computed, but word frequency is retained here as its effect is stronger. Logged frequency values were used as they provide a better fit than raw values.

### 2.2. Statistical analyses

Four mixed models were fitted on the data (Table 1). Models I and II are based on all relevant monosyllabic words in the corpus. Models III and IV were fitted on smaller datasets composed of kinship nouns and verbs that have homophonous function words.

**Table 1:** Models fitted

Model	Lex. words (nb. syll)	Other cat. (nb. syll)
I. Lexical (L) vs. function (F) words	24941 (L)	31649 (F)
II. Lexical words (L) vs. positional nouns (P)	24941 (L)	928 (P)
III. Kinship nouns (L) vs. pronominal use (K)	1007 (L)	1836 (K)
IV. Verbs (L) vs. gramm. counterparts (G)	1901 (L)	2126 (G)

The list of kinship terms and grammaticalized verbs used in the analyses is given below:

Kinship terms: *anh* ‘older brother’, *ba* ‘father’, *bà* ‘grandmother’, *bã* ‘grandmother,3ps’, *bạn* ‘friend’, *câu* ‘uncle’, *con* ‘child’, *cô* ‘aunt’, *cụ* ‘great-grandparent’, *cha* ‘father’, *chị* ‘older sister’, *chú* ‘uncle’, *dâu* ‘daughter-in-law’, *dì* ‘aunt’, *đượng* ‘uncle’, *em* ‘younger sister’, *má* ‘mother’, *mẹ* ‘mother’, *người* ‘person’, *ông* ‘grandfather’, *ổng* ‘grandfather,3ps’, *thầy* ‘master’, *út* ‘youngest child’  
Grammaticalized verbs: *cho* ‘give/to~for’, *để* ‘put/in order’, *đến* ‘arrive/until’, *đi* ‘go/imp’, *lại* ‘come/again’, *lên* ‘go up/up’, *qua* ‘cross/across’, *ra* ‘go out/out’, *tới* ‘arrive/until’, *theo* ‘follow/according’, *vào* ‘enter/in’, *về* ‘return/about’, *vô* ‘enter/in’, *xuống* ‘go down/down’.

In all models, the dependant variable is the *duration* of syllables. Fixed factors include the *log word frequency* and the *lexical category* of the word to which the syllable belongs, and two binary variables establishing if the syllable *precedes a silent pause* or is *sentence-final*. Random factors include intercepts for *speaker* and *word*, as well as random slopes for *all main fixed effects per speaker*. The only exception is that the random slope for *log word frequency by speaker* was dropped in Models III and IV, as the target words were largely concentrated in the same frequency range.

## 3. RESULTS

### 3.1. Function vs. Lexical words

The results of Model I are given in Table 2. First of all, as shown in Figure 1, there are significant frequency effects. Frequent words are shorter overall than rare ones (*logwordfreq*), but this effect seems to cancel out in function words (*LexCat\*logwordfreq*). Moreover, frequency effects are affected in contradictory ways by different types of boundaries. Before silent pauses, frequent words are longer

overall (*Prepausal\*logwordfreq*), while they tend to be shorter sentence-finally (*Sentfinal\*logwordfreq*).

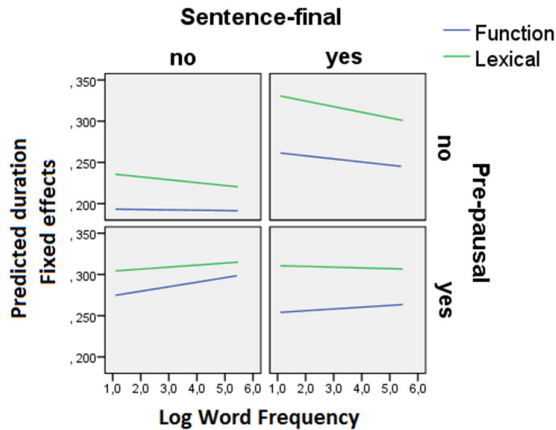
**Table 2:** Model I. Estimates of fixed effects for lexical words (L) vs. function words (F),  $r^2=.35$

Param.	Est	SE	df	t	p
Intercept	.239	.010	24.766	25.06	<.001*
Logwordfreq	-.003	.001	90.38	-3.31	.001*
LexCat=F	-.046	.007	784.10	-6.39	<.001*
Prepausal=Y	.062	.009	46.18	6.75	<.001*
Sentfinal=Y	.099	.010	87.07	9.78	<.001*
LexCat=F*	.003	.002	14534.01	2.027	.043
Logwordfreq					
Prepausal=Y*	.006	.001	42349.95	4.07	<.001*
Logwordfreq					
SentFinal=Y*	-.003	.002	50931.43	-2.03	.042
Logwordfreq					
LexCat=F*	.013	.003	55529.55	4.27	<.001*
Prepausal=Y					
LexCat=F*	-.027	.003	55869.54	-8.04	<.001*
Sentfinal=Y					
Prepausal=Y*	-.089	.003	33109.55	-25.51	<.001*
Sentfinal=Y					

\*  $p < 0.004$  (Bonferroni correction of  $p < 0.05$ )

As can also be seen in Figure 1, there is also an effect of lexical category: all other things being equal, function words are shorter than lexical words (*LexCat*). This effect is even more pronounced sentence-finally (*LexCat\*Sentfinal*). Finally, as expected, there is a clear pre-boundary lengthening. Syllables are 42% longer sentence-finally (*Sentfinal*) and 26% longer before a silent pause. Note that when a syllable is both sentence-final and prepausal, these two types of lengthening do not add up, as indicated by the negative estimate (*Prepausal\*Sentfinal*).

**Figure 1:** Effect of frequency on syllable duration, values predicted by Model 1



### 3.2. Positional nouns

As mentioned in the introduction, positional nouns are not clearly lexical or functional in Vietnamese.

Model II, in Table 3, compares the duration of positional nouns and other lexical words. It turns out that they are not statistically different (*LexCat*). Overall, the other main effects of Model II are similar to Model I in terms of estimates (*Intercept*, *Logwordfreq*, *Prepausal*, *Sentfinal*) even when they are not significant, which is expected as the models are based on largely overlapping data. Interactions tend to have similar coefficients to those of Model I, but with lesser significance. A noticeable difference is that positional nouns are longer than other lexical words in prepausal position (*LexCat\*Prepausal*).

**Table 3:** Model II. Estimates of fixed effects for lexical words (L) vs. positional nouns (P),  $r^2=.36$

Param.	Est	SE	df	t	p
Intercept	.234	.010	20.70	23.27	<.001*
Logwordfreq	-.002	.001	58.51	-1.33	.188
LexCat=P	-.009	.026	6094.85	-.36	.721
Prepausal=Y	.080	.012	258.39	6.85	<.001*
Sentfinal=Y	.063	.015	227.21	4.23	<.001*
LexCat=P*	-.004	.006	10298.17	-.64	.521
Logwordfreq					
Prepausal=Y*	.001	.002	23170.79	.490	.624
Logwordfreq					
SentFinal=Y*	.001	.003	23383.87	.449	.653
Logwordfreq					
LexCat=P*	.063	.015	20642.14	4.32	<.001*
Prepausal=Y					
LexCat=P*	-.041	.017	18284.52	-2.40	.016
Sentfinal=Y					
Prepausal=Y*	-.066	.006	6627.28	-11.53	<.001*
Sentfinal=Y					

\*  $p < 0.004$  (Bonferroni correction of  $p < 0.05$ )

### 3.3. Homophonous words: kinship terms and prepositional verbs

The last two models deal with lexical words that have a grammaticalized counterpart that syntactically behaves like a function word. Model III (kinship terms) is given in Table 4. Here, we see that pronominal kinship terms are not different from corresponding nouns (*LexCat*, *LexCat\*Logwordfreq*, *LexCat\*Sentfinal*), except before silent pauses (*LexCat\*Prepausal*). There is no significant frequency effect in Model III (*Logwordfreq* and its interactions), which is probably caused by the homogeneous frequency distribution of the few words included in it. Finally, there is a trend towards pre-pausal, but not sentence-final, lengthening (*Prepausal*, *Sentfinal*). Once again, these do not add up when a syllable simultaneously precedes a sentence-final silent pause (*Prepausal\*Sentfinal*). The intercept in Model III is low (220 ms), which means that kinship terms (either used as pronouns or nouns) have a duration that is closer to lexical than function words in Model I, but is still a little lower.

**Table 4:** Model III. Estimates of fixed effects for homophonous kinship nouns (L) vs. kinship pronouns (K),  $r^2=.25$

Param.	Est	SE	df	t	p
Intercept	.220	.028	408.70	7.91	<.001*
Logwordfreq	-.003	.006	594.97	-.54	.593
LexCat=K	-.015	.034	449.42	-.43	.670
Prepausal=Y	.156	.070	1672.00	2.23	.026
Sentfinal=Y	.108	.080	1172.65	1.35	.176
LexCat=K*	-.003	.007	539.21	-.38	.705
Logwordfreq					
Prepausal=Y*	-.001	.014	1782.07	-.09	.931
Logwordfreq					
SentFinal=Y*	.006	.016	1423.10	.366	.715
Logwordfreq					
LexCat=K*	-.058	.017	783.88	-3.50	<.001*
Prepausal=Y					
LexCat=K*	-.009	.019	484.14	-.47	.638
Sentfinal=Y					
Prepausal=Y*	-.155	.018	863.51	-8.65	<.001*
Sentfinal=Y					

\*  $p < 0.004$  (Bonferroni correction of  $p < 0.05$ )

Model 4 compares deverbal function words with the verbs from which they seem to have been grammaticalized. Interestingly, these categories do not differ in duration (*LexCat*). Grammatical status does not significantly interact with frequency either (*LexCat\*Logwordfreq*), but it does with preboundary environments, which causes grammaticalized verbs to be lengthened prepausally (*LexCat\*Prepausal*), but shortened sentence-finally (*LexCat\*Sentfinal*).

**Table 5:** Model IV. Estimates of fixed effects for homophonous verbs (L) and grammaticalized words (G),  $r^2=.30$

Param.	Est	SE	df	t	p
Intercept	.251	.023	102.42	10.95	<.001*
Logwordfreq	-.011	.004	1809.61	-2.83	.005
LexCat=G	.029	.026	1660.17	-1.11	.268
Prepausal=Y	-.049	.050	967.87	-.973	.331
Sentfinal=Y	.322	.064	867.96	5.06	<.001*
LexCat=G*	-.001	.006	2117.40	-.193	.847
Logwordfreq					
Prepausal=Y*	.033	.010	3730.28	3.25	.001*
Logwordfreq					
SentFinal=Y*	-.045	.013	2519.52	-3.59	<.001*
Logwordfreq					
LexCat=G*	.048	.015	1741.05	3.26	.001*
Prepausal=Y					
LexCat=G*	-.072	.018	1832.75	-4.00	<.001*
Sentfinal=Y					
Prepausal=Y*	-.093	.018	1850.86	-5.26	<.001*
Sentfinal=Y					

\*  $p < 0.004$  (Bonferroni correction of  $p < 0.05$ )

Frequency effects (*Logwordfreq* and its interactions) are similar to those of Model I, but stronger due to the higher frequency of the words included in Model IV. There is strong sentence-final lengthening (*Sentfinal*),

but no significant main prepausal effect (*Prepausal*). The interaction of the two boundary effects is consistent with what was found in all other models (*Prepausal\*Sentfinal*).

### 3. DISCUSSION

Results from Models I and II show that there is a significant effect of lexical frequency on syllable duration in VN. However, this frequency effect is cancelled out in function words. Yet, function words are still shorter than lexical words. In fact, they are predicted to be 46 ms shorter by Model I, which could be interpreted as a consequence of cliticization.

There are also clear pre-boundary lengthening effects, both before silent pauses and sentence-finally. Prepausal lengthening is more pronounced in function words, while sentence-final lengthening is weaker in this category. We suspect this has to do with the function words that occur in these contexts, but have not been able to isolate a pattern yet.

Durational effects bring new evidence about the status of the VN lexical categories we targeted in the introduction. Positional nouns have durations identical to those of lexical words, which may confirm that they are indeed a class of nouns. Kinship terms and verbs with grammaticalized counterparts show similar behaviours: in both categories, there is no systematic durational difference between lexical words and their grammaticalized counterparts. Although durational effects become murkier in preboundary contexts, this seems to suggest that multifunctional words are not actually split into different lexical entities. Kinship terms are shorter than other lexical words, but this seems attributable to their relatively high frequency.

### 4. CONCLUSIONS

Our results show that high frequency Southern VN lexical words have shorter durations than low frequency words. Function words are not as directly affected by lexical frequency. Evidence of durational differences between function words and lexical words independent of frequency effects suggest that VN may have cliticization processes even in the absence of word stress. In future work, we intend to look at other prosodic cues (intensity,  $f_0$ ) to try to uncover additional evidence of cliticization.

Our corpus also confirms the existence of pre-boundary lengthening in VN. Finally, durational measurements suggest that VN words that have both lexical and grammatical functions (even dramatically different ones) should be analyzed as multifunctional unitary words rather than as independent homophonous words.

## 5. REFERENCES

- [1] Beckman, M., and Edwards, J. 1990. Lengthenings and shortenings and the nature of prosodic constituency, In: J. Kingston, and M. Beckman (eds.). *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*, Cambridge: Cambridge University Press, 152-178.
- [2] Bell, A., Brenier, J., Gregory, M., Girand, C., & Jurafsky, D. 2009. Predictability effects on durations of content and function words in conversational English. *Journal of Memory and Language* 60.1, 92–111.
- [3] Bybee, J. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language variation and change* 14, 261-290.
- [4] Cao, X.H. 1978. Trọng âm và các quan hệ ngữ pháp trong tiếng Việt [Stress and syntactic relations in Vietnamese]. In: *Thông báo ngữ âm học*. Ho Chi Minh City: Social Science Institute.
- [5] Cutler, A., and Butterfield, S. (1990). Syllabic lengthening as a word boundary cue. In Seidl, R. (ed.). *Proceedings of the 3rd Australian International Conference on Speech Science and Technology*, Canberra: Australian Speech Science and Technology Association, 324-328.
- [6] Đỗ-Hurinville D.T. 2010. Les parties du discours en vietnamien: grammaticalisation et catégorialité [Parts of speech in Vietnamese: grammaticalization and categoriality]. *Bulletin de la Société de Linguistique de Paris*. CV-1. 327-370.
- [7] Emeneau, M. 1951. *Studies in Vietnamese (Annamese) Grammar*. Berkeley: University of California Publications in Linguistics
- [8] Fosler-Lussier, E, Morgan, N. 1999. Effect of speaking rates and word predictability on conversational pronunciations, *Speech Communication* 29, 137-158.
- [9] Gahl, S. 2008. Time and Thyme are not homophones: the effect of lemma frequency on word durations in spontaneous speech, *Language* 84.3, 473-496
- [10] Hanske, T. 2013. Serial verbs and caused change of location in Vietnamese, in D. Hole and E. Löbel, eds. *The Linguistics of Vietnamese: an international survey*, Berlin: Walter de Gruyter, 185-214
- [11] Hoàng, T. and Hoàng M. 1975. Remarks on phonological structure of Vietnamese. *Vietnamese Studies* 40. 67-97.
- [12] Jurafsky, D. Bell, M. Gregory and Raymond W. 2001. Probabilistic Relations between Words: Evidence from Reduction in Lexical Production. In Bybee J. and Hopper P. (eds.). *Frequency and the emergence of linguistic structure*. Amsterdam: John Benjamins, 229-254.
- [13] Martini, F. 1958. Notices bibliographiques: langue vietnamienne [Bibliographic note: Vietnamese language]. *Bulletin de l'École française d'Extrême-Orient* 49. 337-348.
- [14] Nespors, M. and Vogel I. 1986. *Prosodic Phonology*. Dordrecht: Foris.
- [15] Ngô, T. N. 1984. *The syllabeme and patterns of word formation in Vietnamese*. PhD Dissertation, New York University.
- [16] Nguyễn, Đ. H. 1997. *Vietnamese: Tiếng Việt không son phấn* [Vietnamese without cosmetics]. Amsterdam/Philadelphia: John Benjamins.
- [17] Pluymaekers, M., Ernestus, M., Baayen, R.H. 2005. Lexical frequency and reduction in spoken Dutch. *Journal of the Acoustical Society of America* 118, 2561-2569.
- [18] Selkirk, E. 2003. The Prosodic Structure of Function Words, In: McCarthy, J. (ed.). *Optimality Theory in Phonology: A Reader*, Blackwell.
- [19] Thompson, L. 1965. *A Vietnamese Grammar*. Seattle: University of Washington Press.
- [20] Turk, A., and Shattuck-Hufnagel, S. 2000. Word-boundary-related duration patterns in English. *Journal of Phonetics* 28, 397-440.