



# Foreign-Language Knowledge Enhances Artificial-Language Segmentation

Annie Tremblay<sup>1</sup>, Mirjam Broersma<sup>2</sup>

<sup>1</sup>University of Kansas, USA

<sup>2</sup>Radboud University, Nijmegen, the Netherlands

atrembla@ku.edu, m.broersma@let.ru.nl

## Abstract

This study investigates whether and how foreign-language knowledge affects the use of non-native cues in speech segmentation. It does so by testing whether Dutch listeners' French knowledge enhances their use of word-final fundamental-frequency (F0) rise—consistent with the typical French prosodic pattern—in artificial-language (AL) speech segmentation. More specifically, this study examines whether Dutch listeners with good French knowledge outperform Dutch listeners with limited French knowledge in the selection of AL words over (nonword or partword) foils, following exposure to an AL with word-final F0 rises. Dutch listeners with good French knowledge completed the AL-segmentation task from Kim et al.'s [2] word-final F0-rise condition. The results were compared to Kim et al.'s [2] Dutch listeners with limited French knowledge and Tremblay et al.'s [1] native French listeners in the same condition. Dutch listeners with good French knowledge performed more accurately than Dutch listeners with limited French knowledge but less accurately than native French listeners on trials with partword foils, with the three groups not differing on trials with nonword foils. Given these results, we propose that foreign-language knowledge can help listeners compute the conditional probability of co-occurrence of successive syllables in an AL and can thus enhance AL speech segmentation.

**Index Terms:** artificial-language speech segmentation, prosody, fundamental frequency, Dutch, French

## 1. Introduction

It is well established that listeners 'hear' unfamiliar languages through the 'ears' of their native (i.e., first) language (L1) [3]. This is also true of speech segmentation: Listeners are better able to segment an unfamiliar language into words when the prosodic cues to word boundaries in the unfamiliar language coincide with those in the L1 than when the unfamiliar language does not contain such cues or contains cues that differ from those in the L1 [1–2, 4–6]. For example, English and Dutch listeners' speech segmentation has been shown to benefit from a fundamental-frequency (F0) rise on the *word-initial* syllable [1–2, 5–6], a finding attributed to the prototypicality of word-initial lexical stress in these languages [7–8]; by contrast, French and Korean listeners' speech segmentation has been shown to be enhanced by a *word-final* F0 rise [1–2, 5, 9], a pattern of results assumed to reflect the importance of intonational phrase-final pitch accents in these two languages [10–11]. The use of F0 cues to word boundaries has thus been deemed to be language-specific [2,5].

What is less clear from existing research, however, is whether listeners' knowledge of a foreign language influences their ability to segment an unfamiliar language when the

unfamiliar language contains cues to word boundaries that are similar to those in the foreign language rather than to those in the L1. Answering these questions can have important implications for the learning of a third language, revealing whether the use of speech segmentation cues can transfer not only from the L1, but also from a subsequently learned language.

The current study sheds light on the role of foreign-language knowledge in speech segmentation by investigating whether and how Dutch listeners' French knowledge influences their use of word-final F0 rise in the segmentation of an unfamiliar language. As indicated above, in Dutch, word-initial syllables tend to have an F0 rise because lexical stress is statistically more likely to be word-initial [8]. By contrast, in French, word-final syllables have an F0 rise in phrase-final position [10–11]. This study seeks to determine whether Dutch listeners' French knowledge will help them segment an unfamiliar language where F0 rise signals word-final boundaries, like it does in French.

An ideal experimental method to answer this question is the Artificial Language (AL) segmentation paradigm [12]. This paradigm consists of two phases: An exposure phase in which participants hear a continuous string of auditory syllables created from a finite set of words (i.e., the AL), and a test phase in which participants decide, in each trial, which of two auditory syllable sequences formed a word in the AL. Studies that use this paradigm typically include a control condition where the only cues to word boundaries in the AL are transitional probabilities: Consecutive syllables that form a word have a high probability of co-occurrence, and syllables that begin or end words have a lower probability of co-occurrence with adjacent syllables that, respectively, precede or follow them. In such a control condition, listeners typically perform above chance in identifying the AL words over foils, indicating that listeners can extract transitional probabilities from the continuous speech stream to locate word boundaries [12–13]. Performance in this control condition can then be compared to performance in experimental conditions where additional cues signal word boundaries in the AL. If listeners' speech segmentation benefits from these additional cues, performance should be higher in these experimental conditions than in the condition that contains only transitional probability cues.

Tremblay et al. [1] used this paradigm to investigate whether listeners' experience with a foreign language would affect their ability to segment an AL. In the AL, F0 cues signaled word-final boundaries. Some participant had French (where F0 also signal word-final boundaries) as their L1 and English (where F0 signals word-initial boundaries) as a foreign language, and other participants had English as L1 and French as a foreign language. Participants were native French listeners with good versus limited knowledge of English and native

English listeners with good versus limited knowledge of French. They completed the AL segmentation task from Kim et al.'s [2] conditions with word-final F0 rises and without F0 cues to word boundaries (control). The forced-choice word-identification task included trials with *nonword* foils (i.e., sequences of AL syllables that never co-occurred in the AL) and trials with *partword* foils (i.e., sequences of two syllables from existing AL words and a third, adjacent syllable from a different, neighboring word in the AL). The results showed that native French listeners with limited English knowledge, but not the English listeners with limited French knowledge, benefited from the presence of F0 cues to word-final boundaries. Furthermore, native English listeners with good French knowledge benefited more from the F0 cues than native English listeners with limited French knowledge. Importantly, English listeners with good French knowledge benefited from F0 cues to word-final boundaries only when the foil was a *nonword*. Tremblay et al. [1] interpreted these results as suggesting that English listeners' French knowledge enabled them to extract the *probability of co-occurrence of two syllables* (nonword foils) but not the *conditional probability of co-occurrence of successive syllables* (partword foils), the latter being more difficult to compute.

Tremblay et al.'s [1] findings suggest that when the unfamiliar language contains cues to word boundaries that differ from those in the L1 but are similar to those in the foreign language, the knowledge of a foreign language modulates listeners' ability to compute the probability of co-occurrence of two syllables in the AL. Crucially, these findings also raise the question of whether the ability to extract the conditional probability of co-occurrence of *successive syllables* in an AL can be enhanced by the knowledge of a foreign language or is determined strictly by whether or not a given word boundary is signaled by the same cue in the L1 and AL.

The current study seeks to shed further light on this question by investigating whether and how Dutch listeners' French knowledge influences their use of word-final F0 rise in the segmentation of an unfamiliar language, thereby revealing whether the findings of Tremblay et al. [1] can be replicated with and/or extended to a different L1. Dutch listeners have been shown to be more reliant on suprasegmental cues to lexical stress compared to English listeners [15–16] because vowel quality plays less of a role in signaling lexical stress in Dutch compared to in English [17]. Given this difference between the two languages, we hypothesize that Dutch listeners' French knowledge may enhance their ability to use word-final F0 cues to extract the conditional probability of co-occurrence of successive syllables in the AL, leading them to outperform Dutch listeners with limited French knowledge on trials with partword foils.

Dutch listeners with good French knowledge completed the AL segmentation task from Kim et al.'s [2] word-final F0-rise condition. Their results were compared to the results of the Dutch listeners with limited French knowledge from Kim et al.'s [2] word-final F0-rise condition [2] and to the results of the native French listeners with limited English knowledge from Tremblay et al. [1] (same condition).

## 2. Method

This study was approved by the human research ethics committees of the University of Kansas and of the Radboud University's Faculty of Arts and Faculty of Philosophy, Theology and Religious Studies.

### 2.1. Participants

Thirty-three native Dutch listeners who were students of French at the bachelor or master level and reported having good French knowledge participated in the study. Participants were screened on the basis of a French proficiency test (cloze test [14]). To ensure that the French proficiency of the Dutch listeners in this study was sufficiently comparable to that of the English listeners with good French knowledge in Tremblay et al. [1], only the Dutch listeners who received a score of 22/45 or higher on the cloze test were included in this study. Two additional participants were excluded because their language background information or cloze test results were missing. The final sample of native Dutch listeners with good French knowledge included 20 participants (mean age: 21.6, standard deviation (SD): 2.9, 14 women). They had first been exposed to French at a mean age of 13.2 years (SD: 2.6), had received on average 8.4 years (SD: 3.2) of French language instruction, and had spent an average of 9.3 months (SD: 11.3) in a French-speaking environment. They received a mean score of 28.8/45 (SD: 4.6) on the cloze test. Note that this is significantly lower than the mean cloze test scores of the native English listeners with good French knowledge in Tremblay et al. [1] (mean: 32.2/45, SD: 5.5;  $t(38) = -2.119, p = .041$ ; matching the two groups in French proficiency while maintaining a minimum sample size of 20 was not possible). All listeners from this group were tested in the Netherlands.

The sample of native Dutch listeners with limited French knowledge from Kim et al. [2] included 20 participants (mean age: 21, SD: 1.3, 14 women; for details, see [2]). These listeners were not students of French. On a scale from 1 ("fluent") to 4 ("poor"), they rated their French as 3.5 (SD: 0.8). These listeners did not know any Korean. All listeners from this group were tested in the Netherlands.

The sample of native French listeners with limited English knowledge from Tremblay et al. [1] included 25 participants (mean age: 21.5, SD: 4.5, 25 women). These listeners reported having limited knowledge of English. All listeners from this group were tested in France.

### 2.2. Materials and Procedures

Participants completed the AL segmentation task from Kim et al.'s [2] word-final F0-rise condition. The task included an exposure phase in which participants listened to the AL and a test phase in which, for each trial, participants decided which of two auditory syllable sequences formed a word in the AL.

The AL consisted of six trisyllabic words ([kapime], [kutəpa], [mupaki], [pətami], [pimatu], [tikəpu]) that did not exist in French or Dutch. The words were produced by a native speaker of Korean (for details, see [2]). All syllables were 252 ms long. The first and second syllables of the words had an F0 of 190 Hz (across the two syllables), and the third syllable had an F0 of 250 Hz (across the entire syllable). Each word was heard a total of 126 times, with no word occurring twice in a row and with no pause between any of the words. There were 20-second fade-in and fade-out periods at the beginning and end of the speech stream. The total duration of the AL was approximately 10 minutes, and participants listened to it twice.

The forced-choice word-identification task contained 36 pairs of trisyllabic sequences, one of which was an AL word and one of which was either a *nonword* foil ([kəpita], [mipaku], [təmuka]) or a *partword* foil ([pupima], [tamiti], [tukapi]). All syllables had an F0 of 190 Hz.

The participants were told that they would be listening to an AL. They were not told anything about French prior to listening to the AL, but they may have been aware of the relevance of French given the recruitment criteria.

### 2.3. Data Analysis

Participants' accuracy was analyzed with logit mixed-effects models using the *lme4* package in R [18], with *p* values calculated from the *lmerTest* package [19]. All models included participant and item as crossed random effects. A first set of models analyzed the data of all three groups, comparing participants' performance against chance and testing for the simple effects of group and foil types and their interaction. A second set of models analyzed only the data of the Dutch listeners with good French knowledge, testing for the simple effects of foil types and French proficiency (cloze test scores) and their interaction.

## 3. Results

Figure 1 shows participants' proportion of correct responses on trials with nonword and partword foils; the error bars represent 95% confidence intervals, and the horizontal line represents chance-level performance.

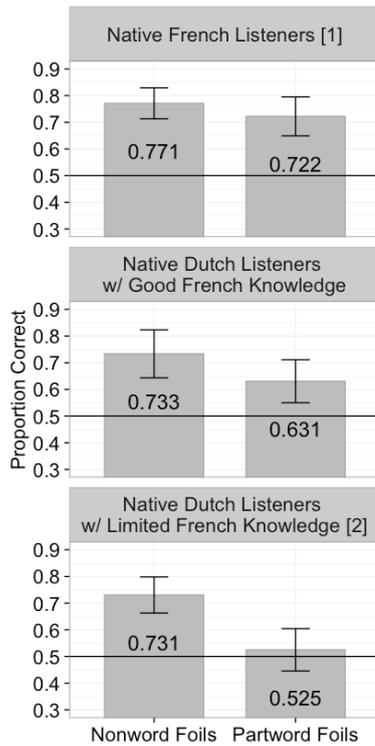


Figure 1: Participants' mean proportion of correct responses; the error bars represent 95% confidence intervals; the horizontal line represents chance-level performance

A first model compared participants' performance on trials with nonword and partword foils against chance. The results of that model, presented in Table 1, show that all groups performed significantly above chance on both foil types except for Dutch listeners with limited French knowledge, who performed above chance only on trials with nonword foils.

Table 1: Logit mixed-effects model comparing participants' performance on trials with nonword and partword foils against chance (est. = estimate, SE = standard error)

Effect	Trial	Est.	SE	<i>z</i>	<i>p</i>
Dutch listeners, good French	Nonword	1.143	0.212	5.400	<.001
	Partword	0.607	0.205	2.959	.003
Dutch listeners, limited French	Nonword	1.098	0.210	5.226	<.001
	Partword	0.113	0.202	< 1	>.1
Native French listeners	Nonword	1.373	0.198	6.947	<.001
	Partword	1.077	0.193	5.583	<.001

A second model tested the fixed effects of group and foil type, and their interaction on participants' accuracy (baseline: Dutch listeners with good French knowledge in trials with nonword foils). This model was compared to a third model testing the fixed effects of group and foil types but not their interaction using the log-likelihood ratio test. The model with the interaction was found to have a better fit than the model without the interaction ( $\chi^2(2) = 9.051, p = .011$ ). The results of the model with the interaction are presented in Table 2. The model did not yield a simple effect of group for Dutch listeners with limited French knowledge or for French listeners, indicating that French knowledge did not affect the groups' performances on trials with nonword foils. The model revealed a simple effect of foil type, indicating that Dutch listeners with good French knowledge performed less accurately on trials with partword foils than on trials with nonword foils. Importantly, the model revealed a marginal interaction between group and foil type for the Dutch listeners with limited French knowledge, suggesting that the effect of group may differ for nonword and partword foils.

Table 2: Logit mixed-effects model testing the effects of group, foil type, and their interaction (est. = estimate, SE = standard error)

Effect	Est.	SE	<i>z</i>	<i>p</i>
Intercept	1.143	0.212	5.400	<.001
Group (Dutch listeners, limited French)	-0.044	0.268	< 1	>.1
Group (French listeners)	0.231	0.258	< 1	>.1
Foil type (partword)	-0.536	0.213	-2.517	.012
Group (Dutch listeners, limited French) × Foil type (partword)	-0.449	0.237	-1.894	.058
Group (French listeners) × Foil type (partword)	0.239	0.234	1.024	>.1

To clarify the meaning of this interaction, the model in Table 2 was relevelled such that partword foils would be the baseline for Dutch listeners with good French knowledge. The relevelled model yielded a marginal simple effect of group for Dutch listeners with limited French knowledge (est. = -0.494, SE = 0.258, *z* = -1.915, *p* = .055), with the latter group performing less accurately than Dutch listeners with good French knowledge on trials with partwords. These results indicate that the marginal group × foil type interaction for Dutch listeners with limited French knowledge comes from the Dutch groups' different performances on trials with partword foils, suggesting that Dutch listeners with good French knowledge outperform Dutch listeners with limited French

knowledge on trials with partwords but not on trials with nonwords.

A fourth model tested the fixed effects of foil type and proficiency score (from which the group mean was subtracted), and their interaction on the accuracy of Dutch listeners with good French knowledge. This model was compared to a fifth model testing the fixed effects of foil type and proficiency score but not their interaction. The model with the interaction was not found to have a better fit ( $\chi^2(1) = 1.019, p > .1$ ). The model that included the fixed effects of foil type and proficiency score was then compared to a sixth model that included only the fixed effect of foil type. The model that included both fixed effects was not found to have a better fit than the model that included only the fixed effect of foil type ( $\chi^2(1) = 1.751, p > .1$ ). These results indicate that French proficiency, as measured by the cloze test [14], was not a significant predictor of the accuracy of Dutch listeners with good French knowledge.

#### 4. Discussion

The present study investigated whether Dutch listeners' French knowledge would enhance their use of F0 rise as a cue to word-final boundaries in AL speech segmentation. Dutch listeners with good French knowledge completed the AL segmentation task from Kim et al.'s [2] word-final F0-rise condition. Their results were compared to the results of the Dutch listeners with limited French knowledge from Kim et al.'s [2] word-final F0-rise condition [2] and to the results of the native French listeners with limited English knowledge from Tremblay et al. [1] (same condition). Dutch listeners with good French knowledge were found to outperform Dutch listeners with limited French knowledge only on trials with partword foils; on trials with nonword foils, all three groups performed similarly. These findings suggest that, with sufficient French knowledge, Dutch listeners *can* extract the conditional probability of co-occurrence of successive syllables (partword foils) after being exposed to an AL where word-final boundaries are signaled by an F0 rise, unlike the findings of Tremblay et al. [1] for the corresponding English listeners. In other words, listeners' ability to extract the conditional probability of co-occurrence of successive syllables in an AL does *not* appear to be determined strictly by whether a given word boundary is signaled by the same cue in the L1 and AL.

A straightforward explanation of the present results is that Dutch listeners may be more sensitive than English listeners to suprasegmental cues such as F0, because these cues play a more important role in signaling lexical stress in Dutch than they do in English [17] (see also [5]). This explanation provides a viable account of Dutch listeners' accuracy on both the trials with partword foils and the trials with nonword foils: On the one hand, the sensitivity to F0 cues may help Dutch listeners with good French knowledge extract the conditional probability of co-occurrence of successive syllables (partword foils), unlike the English listeners with good French knowledge in Tremblay et al. [1], who were *more* proficient in French than the Dutch listeners with good French knowledge in this study; on the other hand, this sensitivity to F0 cues may help Dutch listeners with limited French knowledge extract the probability of co-occurrence of two syllables (nonword foils) from the speech signal even if the F0 cues are used differently in the L1 and AL.

An alternative explanation of the current results exists, however. Although the Dutch listeners in Kim et al. [2] reported having limited French knowledge, all of them had studied French in high school, which was not the case of the English

listeners with limited French knowledge in Tremblay et al. [1]. It is thus possible that the Dutch listeners' limited French knowledge was in fact superior to that of the corresponding English listeners in Tremblay et al. [1], and possibly sufficient to enhance their performance on trials with nonword foils. We favor the first explanation over the second, because French-language instruction in high schools in the Netherlands tends to have a primary focus on reading and writing, making it unlikely that this instruction would result in Dutch listeners' use of F0 cues to word-final boundaries in an AL segmentation task.

The results of this study also showed that French proficiency, as tested by a cloze test [14], was not a significant predictor of the segmentation accuracy of Dutch listeners with good French knowledge. This is unlike the findings of Tremblay et al. [1] for the corresponding English listeners. Note that although the results for this analysis did not reach significance, they trended in the right direction, with segmentation accuracy improving with increasing French proficiency. The current sample of Dutch listeners with good French knowledge did not include as many high-proficiency listeners as the corresponding English listener sample in Tremblay et al. [1]. It is possible that, with more Dutch listeners at a higher French proficiency, a significant effect of proficiency would have been obtained.

The findings of this study have important implications for the learning of a third language. First, they suggest that the use of speech segmentation cues can transfer not only from the L1, but also from a subsequently learned language. Second, they suggest that the functional weight of prosodic cues in the L1 modulates listeners' ability to use the same type of cues in a new language, even if the cues are used differently in the two languages (see also [20]).

One possible limitation of the current study is that the interaction between group and foil type and the simple effect of group for Dutch listeners with limited French knowledge were marginal rather than significant. AL segmentation studies are consistent in showing a great deal of individual variability in listeners' ability to extract transitional probabilities from the signal, even in the absence of additional segmentation cues (e.g., [12]). Hence, within the context of such studies, near significant effects ( $p < .06$ ) are likely to be meaningful.

#### 5. Conclusions

The present study investigated whether and how listeners' knowledge of a foreign language influences their ability to segment an unfamiliar language when the unfamiliar language contains cues to word boundaries that are similar to those in the foreign language rather than to those in the L1. The results showed that the knowledge of a foreign language can positively impact listeners' ability to segment the unfamiliar language into words, enhancing listeners' ability to extract the conditional probability of co-occurrence of successive syllables. Future research should seek to replicate these findings with listeners from a wider variety of language backgrounds.

#### 6. Acknowledgements

This research is based upon work supported by the National Science Foundation under grant no. BCS-1423905 awarded to the first author. Support for this research also comes from a *Language Learning* small research grant awarded to the first author, and a Vidi grant from the Netherlands Organisation for Scientific Research awarded to the second author.

## 7. References

- [1] A. Tremblay, J. Namjoshi, E. Spinelli, M. Broersma, T. Cho, S. Kim, M. T. Martínez-García, and K. Connell, "Experience with a second language affects the use of fundamental frequency in speech segmentation," *PLoS One*, vol. 12, e0181709, 2017.
- [2] S. Kim, M. Broersma, and T. Cho, "The use of prosodic cues in learning new words in an unfamiliar language," *Studies in Second Language Acquisition*, vol. 34, pp. 415–444, 2012.
- [3] A. Cutler, *Native Listening*. Cambridge, MA: MIT Press, 2012.
- [4] M. Ordin, and M. Nespó, "Native language influence in the segmentation of a novel language," *Language Learning and Development*, vol. 12, pp. 461–481, 2016.
- [5] M. D. Tyler and A. Cutler, "Cross-language differences in cue use for speech segmentation," *Journal of the Acoustical Society of America*, vol. 126, pp. 367–376, 2009.
- [6] J. Vroomen, J. Tuomainen, J., and B. de Gelder, "The roles of word stress and vowel harmony in speech segmentation," *Journal of Memory and Language*, vol. 38, pp. 133–149, 1998.
- [7] A. Cutler, and D. M. Carter, "The predominance of strong initial syllables in the English vocabulary," *Computer Speech and Language*, vol. 2, pp. 133–142, 1987.
- [8] R. Schreuder, and R. H. Baayen, "Prefix stripping re-visited," *Journal of Memory and Language*, vol. 33, pp. 357–375, 1994.
- [9] O. Bagou, C. Fougéron, and U. Frauenfelder, "Contribution of prosody to the segmentation and storage of "words" in the acquisition of a new mini-language," *Proceedings of the Speech Prosody 2002 Conference*, pp. 159–162, 2002.
- [10] S.-A. Jun, "The Accentual Phrase in the Korean prosodic hierarchy," *Phonology*, vol. 15, pp. 189–226, 1998.
- [11] S.-A. Jun and C. Fougéron, "Realizations of accentual phrase in French intonation," *Probus*, vol. 14, pp. 147–172, 2002.
- [12] J. R. Saffran, E. L. Newport, and R. N. Aslin, "Word segmentation: The role of distributional cues," *Journal of Memory and Language*, vol. 35, pp. 606–621, 1996.
- [13] J. R. Saffran, "The use of predictive dependencies in language learning," *Journal of Memory and Language*, vol. 44, pp. 493–515, 2001.
- [14] A. Tremblay, "Proficiency assessment standards in second language acquisition research," *Studies in Second Language Acquisition*, vol. 33, pp. 339–372, 2011.
- [15] N. Cooper, A. Cutler, and R. Wales, "Constraints of lexical stress on lexical access in English: Evidence from native and non-native listeners," *Language and Speech*, vol. 45, pp. 207–228, 2002.
- [16] V. J. van Heuven, and M. de Jonge, "Spectral and temporal reduction as stress cues in Dutch," *Phonetica*, vol. 68, pp. 120–132, 2011.
- [17] A. M. C. Sluijter, and V. J. van Heuven, "Acoustic correlates of linguistic stress and accent in Dutch and American English," *Proceedings of the International Congress of Spoken Language Processing*, 1996.
- [18] D. Bates, B. Maechler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, pp. 1–48, 2015.
- [19] A. Kuznetsova, B. Brockhoff, and H. Christensen, "Tests in linear mixed effects models," retrieved from <https://cran.r-project.org/web/packages/lmerTest/index.html>, 2016.
- [20] A. Tremblay, M. Broersma, and C. E. Coughlin, "The functional weight of a prosodic cue in the native language predicts the learning of speech segmentation in a second language," *Bilingualism: Language and Cognition*, vol. 21, pp. 640–652, 2018.