JKSCI

한국컴퓨터정보학회논문지
Journal of The Korea Society of Computer and Information
Vol. 31 No. 1, pp. 359-367, January 2026
https://doi.org/10.9708/jksci.2026.31.01.359

Acoustic Characteristics of English Interjections and Discourse Markers: Focused on Speech in Influencers' Vlogs

Namjin Lee*, Tae-Yeoub Jang**

*Student, Dept. of English Linguistics, Hankuk University of Foreign Studies, Seoul, Korea
**Professor, Dept. of English Linguistics and Language Technology, Hankuk University of Foreign Studies, Seoul, Korea

[Abstract]

This study investigates the prosodic realization of pragmatic markers in spontaneous English vlog speech, focusing on interjections, discourse markers, and hybrid forms. A total of 250 tokens (86 items) extracted from English-language YouTube vlogs were analyzed using acoustic–phonetic measures of speech rate, pitch, intensity, and pause behavior. The results show that speech rate is the most robust distinguishing cue: interjections are produced more slowly and occur more frequently with pauses than discourse markers and hybrid forms. Pitch and intensity do not reliably differentiate categories in absolute terms, although interjections exhibit localized pitch elevation and increased loudness relative to surrounding speech. By moving beyond predominantly impressionistic classifications, this study demonstrates the feasibility of quantitatively distinguishing pragmatic marker types and highlights temporal organization as central to pragmatic prosody, with implications for automatic discourse analysis and data-driven speech-based learning systems.

▸ **Key words**: Pragmatic markers, Interjections, Discourse markers, Prosody, Acoustic analysis, Vlogs

[요 약]

본 연구는 자연 발생적 영어 발화에서 사용되는 화용 표지의 운율적 실현 양상을 분석하여 화용 기능과 음성적 특성 간의 관계를 규명하는 것을 목적으로 한다. 이를 위해 영어 유튜브 브이로그 10편을 대상으로 수작업을 통해 분석하여 총 250개의 화용 표지 토큰을 추출한 후, 발화 속도, 기본주파수, 강도, 휴지 양상 등의 음향음성학적 특징을, 토큰 자체 및 주변 발화와의 대비 수준에서 정량적인 지표를 추출하여 통계적인 분석을 진행하였다. 그 결과, 간투사는 담화 표지 및 혼합 유형에 비해 유의미하게 느린 발화 속도와 높은 휴지 동반 빈도를 보였다. 반면, 음높이와 강도는 유형 간 뚜렷한 차이를 보이지 않았다. 본 연구는 전통적으로 연구자의 주관적 판단에 의존했던 화용 표지의 분류를 정량적인 지표를 통해 진행할 수 있음을 증명하였고, 추후 이런 특징들을 사용하여 자동 음성처리 시스템 성능 향상을 위한 직접적인 응용이 가능하리라고 기대된다.

▸ **주제어**: 화용 표지, 간투사, 담화표지, 운율, 음향 분석, 유튜브 브이로그

• First Author: Namjin Lee, Corresponding Author: Tae-Yeoub Jang
*Namjin Lee (einstein@hufs.ac.kr), Dept. of English Linguistics, Hankuk University of Foreign Studies
**Tae-Yeoub Jang (tae@hufs.ac.kr), Dept. of English Linguistics and Language Technology, Hankuk University of Foreign Studies
• Received: 2025. 12. 22, Revised: 2026. 01. 02, Accepted: 2026. 01. 19.

http://www.ksci.re.kr pISSN:1598-849X | eISSN:2383-9945

I. Introduction

Spoken language research has long recognized that pragmatic meaning is closely associated with prosodic form. While earlier approaches often assumed that pragmatic functions are signaled through relatively stable acoustic cues such as pitch or intensity, recent experimental phonetic research with interactional data suggests that pragmatic meaning is more commonly conveyed through temporal alignment, rhythmic organization, and coordination with surrounding speech [1], [2]. From this perspective, prosody functions not merely as an expressive overlay but as a relational mechanism through which pragmatic meaning is constructed in context.

Pragmatic markers, including interjections and discourse markers, play a central role in organizing discourse, expressing speaker stance, and guiding interpretation in real time [3]-[5]. Although syntactically peripheral and often optional, their meanings are highly context dependent and emerge through sequential placement and prosodic realization rather than through lexical semantics alone [6], [7]. Prosody is therefore particularly central to their interpretation, with timing, prominence, and pause behavior shaping whether a form is understood as affective, discourse-organizing, or multifunctional [2], [8].

Despite this recognized importance, systematic acoustic research on pragmatic markers remains limited. Much of the existing literature has focused on functional classification or discourse distribution, while prosodic properties are frequently treated descriptively rather than as targets of quantitative analysis [4], [5]. As a result, it remains unclear whether distinctions among pragmatic marker types correspond to invariant acoustic cues or are better understood as gradient differences in temporal and interactional integration.

This issue is especially salient for multifunctional forms that shift between interjective and discourse-marking uses depending on context [9], [10]. Continuum-based approaches to pragmatic meaning therefore predict that such functional variability should be reflected in degrees of prosodic prominence and integration rather than in discrete categorical contrasts [11], [12]. Empirical acoustic evidence directly evaluating this prediction, however, remains relatively scarce.

Spontaneous speech in digitally mediated contexts such as YouTube vlogs provides an effective testing ground for these claims. In this genre, pragmatic markers occur frequently and are deployed in extended stretches of naturally occurring talk, making prosodic resources particularly salient [13], [14]. The present study investigates the prosodic realization of pragmatic markers in spontaneous English vlog speech, focusing on interjections, discourse markers, and their hybrid forms. Using acoustic-phonetic analysis, it examines speech rate, pitch, intensity, and pause behavior relative to surrounding utterances in order to clarify how pragmatic distinctions are realized prosodically in natural speech.

Previous studies employing computational voice analysis on relatively large-scale speech datasets have primarily focused on the automated classification of pragmatic markers. However, their accuracy has likely been constrained by the absence of appropriately pre-defined features. In contrast, the present study adopts a discourse-pragmatic, function-first approach, combining fine-grained manual annotation with acoustic-phonetic analysis to examine how pragmatic distinctions are prosodically encoded in naturally occurring speech.

The remainder of this paper is structured as follows. Section II reviews previous research on pragmatic markers and prosody, establishing the theoretical background of the study. Section III describes the data, annotation procedures, acoustic measures, and statistical methods employed in the analysis. Section IV presents the results and

discusses their implications for relational models of pragmatic prosody and pragmaticalization. Section V presents the concluding remarks and outlines future directions.

II. Background: Pragmatic markers, prosody, and functional continua

Pragmatic markers constitute a heterogeneous class of expressions that operate primarily at the interpersonal and discourse-organizational levels of spoken interaction. Rather than contributing propositional content, they guide interpretation through sequential placement and prosodic realization, and are therefore commonly characterized as procedural elements whose meaning emerges dynamically in interaction [3], [4], [6].

Within this broad category, interjections and discourse markers have often been treated as analytically distinct types. Interjections are typically described as affective or reactive expressions that convey immediate emotional or epistemic responses and are frequently produced with prosodic salience, including temporal expansion and enhanced prominence [11], [15]–[17]. Discourse markers, by contrast, function primarily to organize discourse and manage interactional flow, signaling transitions, coherence relations, or speaker alignment. Prosodically, they tend to be rhythmically integrated into surrounding speech and display reduced prominence, reflecting their role in maintaining discourse continuity [3], [5], [8], [12].

Between these prototypical categories, a range of multifunctional or hybrid forms has been identified. Expressions such as 'okay,' 'listen,' and 'bro' may simultaneously convey speaker stance and contribute to discourse organization, with their interpretation depending on sequential context and prosodic delivery [9], [18]. Such observations have motivated continuum-based approaches to pragmatic meaning, which conceptualize pragmatic functions as gradient rather than discrete and predict corresponding differences in degrees of prosodic prominence and integration.

Continuum-based accounts are closely related to theories of pragmaticalization, which describe diachronic shifts whereby expressions move from expressive or lexical uses toward increasingly discourse-organizing functions [18], [19]. These accounts propose that such functional shifts are accompanied not only by semantic bleaching and structural integration but also by systematic changes in phonetic realization, including reduced duration, smoother temporal alignment, and diminished prominence. Despite the prominence of these claims in theoretical work, empirical acoustic evidence directly linking pragmatic function to consistent prosodic patterns remains relatively limited.

From an interactional perspective, prosody—particularly timing, rhythm, and pause placement—plays a central role in organizing real-time speech [2], [21]. Early phonetic work already emphasized temporal organization as a foundational property of spoken language rather than a secondary expressive layer [20]. Taken together, previous studies suggest that pragmatic marker distinctions are best understood in terms of relative prosodic and interactional integration within the speech stream, a hypothesis that can be systematically evaluated through acoustic-phonetic analysis.

III. Data and Methodology

1. Data

The dataset consists of 250 pragmatic marker tokens manually extracted from publicly available English-language YouTube vlogs. Excluding duplicate lexical forms, these tokens correspond to 86 distinct lexical or phrasal items. The data were curated specifically for this study to capture spontaneous speech in a contemporary digital context, in which pragmatic markers are frequent and functionally diverse.

The sample includes ten speakers, evenly divided by nationality, with five speakers from the United States and five from the United Kingdom (See Appendix A for speaker details). Variation in gender and racial background is also represented. All speakers are high-visibility content creators with subscriber counts exceeding one million, which helps reduce the likelihood that observed prosodic patterns are attributable to idiosyncratic speaker characteristics rather than pragmatic function. Figure 1 presents a word cloud showing the frequency distribution of pragmatic marker tokens, highlighting the most common expressions. A comprehensive list of target pragmatic markers is presented in Appendix B.

Fig. 1. Wordcloud of Pragmatic Markers' Frequency

2. Annotation and Categorization

Pragmatic markers were categorized based on functional role and prosodic behavior rather than lexical form alone. Classification was informed by the interaction of pragmatic function, prosodic realization, and sequential context.

Interjections were defined as affective or reactive expressions characterized by prosodic prominence, including temporal expansion, elevated pitch or intensity, and a higher likelihood of adjacency to pauses. Discourse markers were defined as expressions whose primary function is to organize discourse and manage interactional flow; these forms typically exhibit rhythmic integration into surrounding speech and reduced prosodic prominence. Hybrid markers were defined as multifunctional forms that combine stance-related and discourse-organizing functions and often display mixed prosodic profiles. This operationalization reflects the gradient and context-sensitive nature of pragmatic marker use.

For example, the form 'okay' was classified as a discourse marker when produced with reduced duration and smooth rhythmic integration into the surrounding utterance, signaling transition or continuation. In contrast, tokens of 'okay' accompanied by temporal expansion, pitch prominence, and adjacency to pauses were classified as interjection. Tokens exhibiting mixed prosodic properties were categorized as hybrid.

3. Acoustic Measures

Acoustic analysis was conducted using Praat [22]. For each pragmatic marker token, measurements were extracted at both the token level and relative to the surrounding utterance in order to capture relational prosodic behavior. Acoustic dimensions included speech rate (in syllables per second, syl/sec), mean fundamental frequency (F0, in Hz), and mean intensity (in dB).

Pause behavior was also analyzed by identifying silent intervals of 200 ms or longer occurring immediately before and after each token. Both pause frequency (N) and pause duration (in millisecond) were recorded. Figure 2 presents a segmentation and measurement example, with Appendix C providing illustrations of the three marker types compared. Table 1 summarizes the dataset structure and the acoustic variables analyzed.

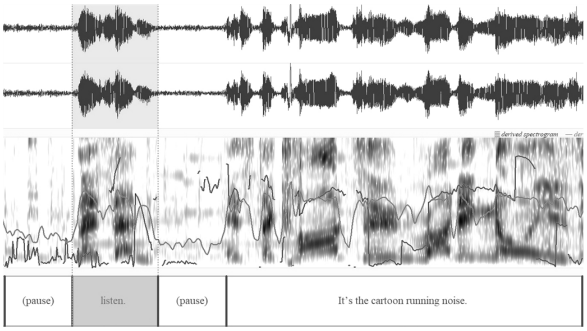

Fig. 2. Acoustic Measurement of the target "listen" in hybrid usage

Table 1. Dataset Structure

Variable	Definition / Unit
Speaker	Source identifiers
Nationality	Speaker's nationality
Gender	Speaker's gender
Target	Analyzed token
Type	Interjection (i), Discourse Marker (d), Hybrid (h)
Rate_t, Rate_u (syll/s)	Local and utterance-level speech rate
F0_t, F0_u (Hz)	Mean F0 of target & utterance
Int_t, Int_u (dB)	Mean intensity of target & utterance
Npau	Number of pauses immediately adjacent to the token
pau1 (ms)	Duration of the pause preceding the token, defined as a silent interval of at least 200 milliseconds
pau2 (ms)	Duration of the pause following the token, also defined as a silent interval of at least 200 milliseconds
pau_total (ms)	Combined duration of pre- and post-token pauses

Speech rate was measured at the token level (Rate_t) and utterance level (Rate_u) to assess whether pragmatic markers are produced significantly faster or slower than the utterances containing them. Fundamental frequency and intensity were likewise extracted for the target token (F0_t, Int_t) and the surrounding utterance (F0_u, Int_u), enabling relational comparison across pragmatic marker types. Pause behavior was characterized in terms of both frequency (Npau) and duration, with silent intervals occurring before a token (pau1) or after a token (pau2) classified as pauses when their duration was at least 200 ms.

4. Statistical Analysis

Statistical analyses examined differences in prosodic realization across pragmatic marker types and between pragmatic markers and their surrounding utterances. One-way analyses of variance (ANOVAs) were conducted to assess overall differences among interjections, discourse markers, and hybrid forms. Where appropriate, post hoc comparisons were performed using Tukey's HSD tests. Targeted paired-samples t-tests were used for theoretically motivated contrasts between token-level and utterance-level measures, and effect sizes were computed using Cohen's d. Assumptions of parametric testing were evaluated using Shapiro-Wilk tests.

IV. Results

The acoustic measurement results for each feature across the three types of pragmatic markers are summarized in Table 2.

Table 2. Summary of Overall Mean Acoustic Measures for Pragmatic Markers by Type

Feature	Type	N	Mean	SD	F	p / Sig.
Speech Rate (syl/sec)	Interjection	92	3.23	1.65	20.79	<.001 ***
	Discourse Marker	116	5.13	2.37		
	Hybrid	42	4.51	2.33		
F0 (Hz)	Interjection	92	205.44	90.36	0.86	.423 (n.s.)
	Discourse Marker	116	192.60	50.95		
	Hybrid	42	194.39	77.25		
Intensity (dB)	Interjection	92	75.15	4.16	0.09	.914 (n.s.)
	Discourse Marker	116	74.88	4.52		
	Hybrid	42	74.96	5.65		

Across analyses, speech rate emerged as the most robust cue distinguishing pragmatic marker types. Interjections were produced significantly more slowly and were more frequently adjacent to pauses than discourse markers and hybrid forms, indicating systematic temporal expansion. Discourse markers and hybrid forms, by contrast, exhibited comparable articulation rates and strong rhythmic integration with surrounding speech, consistent with their role in maintaining discourse continuity. Pitch and intensity did not reliably differentiate pragmatic marker types in absolute terms.

However, the comparison between the target pragmatic marker and its surrounding utterances, as presented in Table 3, revealed more selective patterns: interjections and discourse markers showed modest pitch elevation relative to surrounding speech, while only interjections displayed consistent intensity enhancement. Hybrid forms showed no systematic deviation across measured dimensions, reflecting a high degree of

prosodic incorporation.

Table 3. Differences in Mean Acoustic Features between Target Pragmatic Markers (by type) and Surrounding Utterances

Feature	Type	Target	Utterance	*t*	p / Sig.
Speech Rate (syl/sec)	Interjection	3.23	4.94	−7.69	<.001 ***
	Discourse Marker	5.13	5.11	0.11	.910 (n.s.)
	Hybrid	4.51	4.64	−0.38	.707 (n.s.)
F0 (Hz)	Interjection	205.44	183.35	2.57	.012*
	Discourse Marker	192.60	184.80	2.22	.028*
	Hybrid	194.39	188.64	0.44	.659 (n.s.)
Intensity (dB)	Interjection	75.15	73.94	3.40	.001**
	Discourse Marker	74.88	74.62	0.96	.338 (n.s.)
	Hybrid	74.96	74.57	0.65	.522 (n.s.)

Pause behavior, summarized in Table 4, further underscored differences in interactional integration. Interjections occurred more frequently adjacent to pauses and were associated with longer pre-token pauses, highlighting the role of pausing as an interactional resource for marking affective salience and discourse boundaries. In contrast, discourse markers and hybrid forms showed minimal deviation in pause behavior, aligning with their discourse-organizing function.

Table 4. One-way ANOVA Results for Pause-related Features (Npau, pau1, pau2, pau_total)

Source	Types	N	Sum	Mean	SD	F	p / Sig.
Pause Frequency							
Npau (n)	*i*	92	63	0.68	0.82	8.88	< .001***
	d	116	34	0.29	0.54		
	h	42	18	0.43	0.59		
Pre-pause Duration							
pau1 (ms)	*i*	92	39864	433.30	1493.85	3.73	0.025*
	d	116	9309	80.25	199.44		
	h	42	6945	165.36	487.25		
Post-pause Duration							
pau2 (ms)	*i*	92	15700	170.65	367.15	2.92	0.056
	d	116	8903	76.75	232.14		
	h	42	3845	91.55	206.55		
Total Pause Duration							
pau_total (ms)	*i*	92	55564	603.96	1643.16	4.79	0.009*
	d	116	18212	157.00	349.93		
	h	42	10790	256.90	566.36		

Figure 3 provides a schematized visualization of the acoustic features of target pragmatic markers, with speech rate, pitch, and intensity represented by size, color, and darkness, respectively.

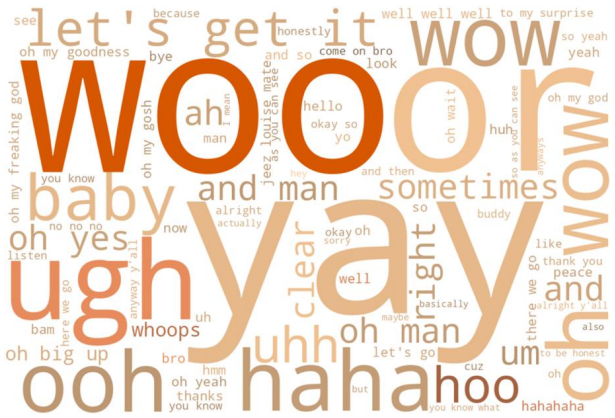

Fig. 3. Word cloud of pragmatic markers with size reflecting speech rate, color representing pitch, and intensity represented by darkness

V. Conclusion

Taken together, these findings reveal a hierarchy of prosodic marking in which interjections are maximally marked, discourse markers exhibit limited and targeted prominence, and hybrid forms resist stable categorization. Rather than reflecting invariant acoustic signatures, pragmatic marker distinctions are encoded through gradient differences in temporal and interactional integration. These results support relational models of pragmatic prosody and align with theories of pragmaticalization that predict increased temporal integration as expressions shift toward discourse-organizing functions.

It should be noted that vlog speech may be subject to editing practices such as cuts, jump edits, or post-production processing. However, the pragmatic markers examined in this study typically occur within locally produced stretches of speech, where relative timing and prosodic integration remain interpretable despite potential editing. Future research may further examine how different

production styles and degrees of post-production systematically shape pragmatic prosody across digital speech genres.

While the focus on vlog speech may limit generalizability, future research can extend this approach to other discourse settings, genres, and languages. By foregrounding timing, rhythm, and alignment, this study provides empirically grounded constraints for prosody-aware approaches to automatic discourse analysis and speech-based language technologies.

ACKNOWLEDGEMENT

This work was supported by Hankuk University of Foreign Studies Research Fund of 2025.

REFERENCES

[1] T. Johnstone and K. R. Scherer, "Vocal Communication of Emotion," in Handbook of Emotions, 2nd ed., M. Lewis and J. M. Haviland-Jones, Eds., pp. 220-235, Guilford Press, 2000.

[2] J. Local and G. Walker, "Methodological Imperatives for Investigating the Phonetic Organisation and Phonological Structures of Spontaneous Speech," Phonetica, Vol. 62, No. 2-4, pp. 120-130, 2005. DOI: 10.1159/000090093

[3] D. Schiffrin, "*Discourse Markers*" Cambridge University Press, 1987.

[4] B. Fraser, "What Are Discourse Markers?" Journal of Pragmatics, Vol. 31, No. 7, pp. 931-952, 1999. DOI: 10.1016/S0378-2166(98)00101-5

[5] K. Aijmer, "*English Discourse Particles: Evidence from a Corpus*" John Benjamins, 2002.

[6] J. Verschueren, "*Understanding Pragmatics*" Oxford University Press, 1999.

[7] E. Goffman, "*Forms of Talk*" University of Pennsylvania Press, 1981.

[8] M. Swerts, "Filled Pauses as Markers of Discourse Structure," Journal of Pragmatics, Vol. 30, No. 4, pp. 485-496, 1998. DOI: 10.1016/S0378-2166(98)00014-9

[9] N. R. Norrick, "Pragmatic Markers, Interjections and Discourse," Catalan Journal of Linguistics, Vol. 6, pp. 159-168, 2007.

[10] J. Cheshire, "Discourse Variation, Grammaticalisation and Stuff Like That," Language in Society, Vol. 36, No. 2, pp. 239-259, 2007.

[11] F. K. Ameka, "Interjections: The Universal Yet Neglected Part of Speech," Journal of Pragmatics, Vol. 18, No. 2-3, pp. 101-118, 1992. DOI: 10.1016/0378-2166(92)90048-G

[12] L. J. Brinton, "*The Comment Clause in English*" Cambridge University Press, 2008.

[13] M. Frobenius, "Audience Design in Monologues: How Vloggers Involve Their Viewers," Journal of Pragmatics, Vol. 72, pp. 59-72, 2014. DOI: 10.1016/j.pragma.2014.02.008

[14] J. Androutsopoulos and J. Tereick, "YouTube: Language and Discourse Practices in Participatory Culture," in *The Routledge Handbook of Language and Digital Communication*, pp. 354-370, Routledge, 2016.

[15] M. Dingemanse, "Interjections," in *The Oxford Handbook of Word Classes*, pp. 411-438, Oxford University Press, 2023.

[16] D. Wilkins, "Interjections as Deictics," Journal of Pragmatics, Vol. 18, No. 2-3, pp. 119-158, 1992. DOI: 10.1016/0378-2166(92)90049-H

[17] C. Goddard, "Interjections and Emotion," Emotion Review, Vol. 6, No. 1, pp. 53-63, 2014.

[18] E. C. Traugott and R. B. Dasher, "*Regularity in Semantic Change*" Cambridge University Press, 2002.

[19] P. J. Hopper and E. C. Traugott, "*Grammaticalization*" Cambridge University Press, 2003.

[20] D. Abercrombie, "*Elements of General Phonetics*" Edinburgh University Press, 1967.

[21] G. Walker, "Coordination and Interpretation of Vocal and Visible Resources," Language and Speech, Vol. 55, No. 1, pp. 141-163, 2012. DOI: 10.1177/0023830911428858

[22] P. Boersma and D. Weenink, "Praat: Doing Phonetics by Computer," Computer Program, Version 6.3, 2023.

Appendix A. Speaker Profile Summary

N	Creator	Nation	Sex	Race	Subscribers	Sample Video (Title + Date)
01	Casey Neistat	US	M	White	12.6M	Planned Obsolescence and the Death of Great Products (11 Jul 2025)
02	Emma Chamberlain	US	F	White	12.1M	Facing My Fear of Flying (Alone) (1 Sep 2025)
03	BenDeen	US	M	East Asian	1.49M	The Spiciest Ramen in the World (14 Apr 2022)
04	Lilly Singh	US	F	South Asian	14.2M	Releasing My First Movie (It's a Sexy Comedy!) (26 Sep 2025)
05	Logan Paul	US	M	White	23.6M	My Daughter's 1st Birthday SURPRISE emotional (2 Oct 2025)
06	Zoe Sugg	UK	F	White	5.0M	Autumn Porch, Sunday Roast & Halloween Haul (19 Oct 2025)
07	Niko Omilana	UK	M	Black	8.18M	I Pranked America's Most Racist Man (14 May 2022)
08	Yung Filly	UK	M	Black	1.8M	THE ALPHABET GAME (5 Oct 2024)
09	Patricia Bright	UK	F	Black	2.71M	The TRUTH About Where I've Been··· (5 May 2025)
10	Ali-A	UK	M	White	19.7M	*Fortnite's NEW Fortnitemares UPDATE is INSANE!* (9 Oct 2025)

Appendix B. All Target words by Types

Type	Targets (Duplicates Removed)
Interjection (N= 92)	*ah, baby, bam, bro, buddy, bye, haha, hahahaha, hello, hey, hmm, hoo, huh, jeez louise mate, no no no, oh, oh my god, oh my goodness, oh my gosh, oh wow, okay, ooh, oops, peace, sorry, thank you, thanks, ugh, uh, uhh, um, whoops, woo, wow, yay, yeah.*
Discourse Marker (N=116)	*I mean, actually, alright, alright y'all, also, and, and listen, and so, and then, anyways, anyway y'all, as you can see, basically, because, bro, but, come on, like, listen, look, no, now, okay, okay so, or, right, see, so, so as you can see, sometimes, well, yeah, you know, you know what.*
Hybrid (N=42)	*and man, big up, bro, clear, come on, come on bro, cuz, here we go, hey, honestly, let's get it, let's go, listen, man, maybe, oh big up, oh man, oh yeah, oh yes, oh wait, so yeah, to be honest, to my surprise, there we go, um, well well well, yeah, yo.*

Appendix C. Example Waveforms and Spectrograms

Figure	Token	Category	Key prosodic characteristics illustrated
C1	okay	Interjection	Temporal expansion; pause adjacency; reduced rhythmic integration
C2	listen	Hybrid	Mixed prominence; partial rhythmic integration; transitional function
C3	listen	Discourse marker	Rhythmic integration into following clause; minimal temporal expansion

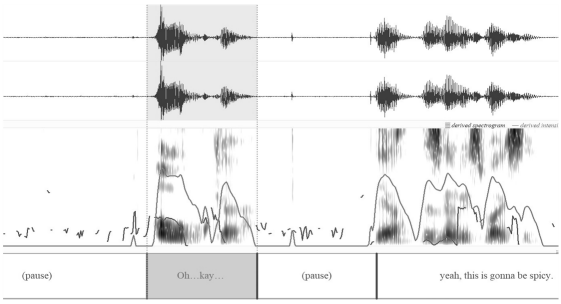

Fig. C1. Waveform and spectrogram of 'okay' produced as an interjection. The token exhibits temporal expansion and clear adjacency to surrounding pauses, resulting in reduced rhythmic integration with the following utterance.

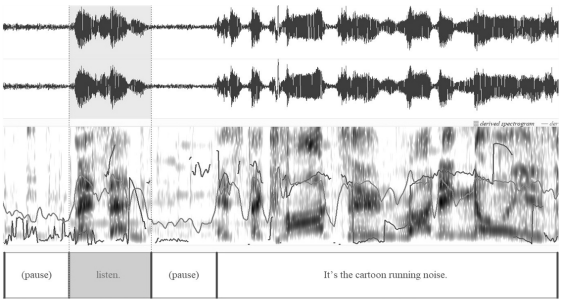

Fig. C2. Waveform and spectrogram of 'listen' exhibiting hybrid pragmatic use. The token displays partial temporal expansion and increased prominence relative to surrounding speech, while simultaneously functioning as a transition into the subsequent utterance.

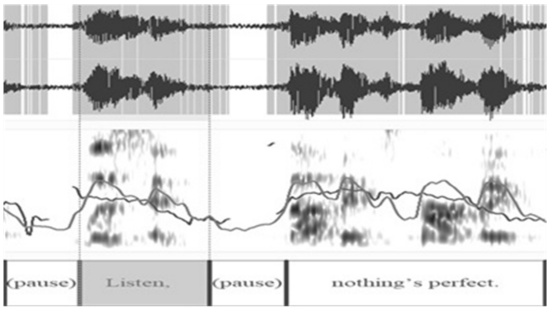

Fig. C3. Waveform and spectrogram of 'listen' produced as a discourse marker. Although a brief pause precedes the token, it is produced with minimal temporal expansion and is rhythmically integrated into the following clause, reflecting its discourse-organizing function.

Authors

Namjin Lee received the B.A. degree in Journalism from Hankuk University of Foreign Studies, Seoul, South Korea. He received the M.A. degree in Interpretation and Translation (Korean-English) from Chung-Ang University, Seoul, South Korea. He is currently pursuing the Ph.D. degree in Linguistics, with a specialization in Phonetics, at Hankuk University of Foreign Studies. His research interests include phonetic analysis, speech production and perception, and Korean-English linguistic comparison.

Tae-Yeoub Jang received the B.S. and M.S. degrees in English Linguistics from Hankuk University of Foreign Studies, Korea, in 1987 and 1989, respectively. and the Ph.D. degree in Automatic Speech Recognition from University of Edinburgh, United Kingdom. Dr. Jang joined the faculty of the Department of English Linguistics at Hankuk University of Foreign Studies, Seoul, Korea, in 2002. He is currently a Professor in the Department of English Linguistics and Language Technology, Hankuk University of Foreign Studies. He is interested in spoken language processing, automatic speech recognition and experimental phonetics.