

# レビューの新たな品質メトリクス収集 「エーオース(Eōs)法」の提案

～レビュー会議の会話データを分析し、発言に関する心理的パラメータを測定～

第38年度（2022年度）ソフトウェア品質管理研究会  
研究コースⅡ ソフトウェアレビュー

「TEAM-ASTRAIÄ（チーム アストライアー）」

研究員：星野 智彦（株式会社アイソフ）  
村上 薫（株式会社日立システムズ）  
芳沢 圭一（株式会社オーグス総研）  
主査：中谷 一樹（TIS株式会社）  
副主査：上田 裕之（株式会社DTSインサイト）  
アドバイザー：安達 賢二（株式会社HBA）



レビューの質を上げたい！



レビューの質を上げたい!

いろいろ取り組んでいるが...



何かもっと本質的で大事なことが?



参加者全員で協力できていますか？



参加者全員で協力できていますか？

発言できていない人も…



発言しやすい雰囲気にしよう！

参加者全員が**発言しやすい場**にするために



人の心理に切り込む！

**レビュー-会議の雰囲気**について考えてみましょう！

でも雰囲気って言葉で表現するのが難しい…

どうしたら伝わるかな？



「ゆっくり」の動画  
がいいかも！



雰囲気異なる2つのレビュー会議を  
「ゆっくり」動画でご覧ください

※「ゆっくり」動画の利用許可を得ています！



# ① フラックレビュー

# ② ホワイトレビュー



でもまあ、ほんと喋りにくそうな場だったよね



※動画から抜粋



すごく会話も弾んでたんだぜ!





どちらの雰囲気気でレビューをやりたいですか？

どちらの雰囲気てレビューをやりたいですか？



もちろん  
ホワイトっしょ！

発言しやすい場

# 仮説

発言しやすい場

全員が建設的に  
発言

レビュー効果の  
向上

ウォーターフォール開発のレビューでは、  
従来、指摘する側とされる側という関係でレビューを実施してきた

全員で品質を上げることを考えるような場を実現することでレビューの効果が上がるはず

# 仮説

発言しやすい場

全員が建設的に  
発言

レビュー効果の  
向上



研究対象



発言しやすい場とは？

発言しやすい場とは？

 注意!!!

「和んだ雰囲気 = 発言しやすい場」ではない!



和気あいあいの感じだが…  
お互い遠慮している



喧嘩をしているように見えて…  
真剣に発言している

発言しやすい場とは？



発言しにくい場

改善する



発言しやすい場

測れないものは改善できない!!!

発言しやすさを客観的に測るメトリクスを探究する!!!



# レビューの新たな品質メトリクス収集 「エーオース( $\bar{E}\bar{o}s$ )法」の提案

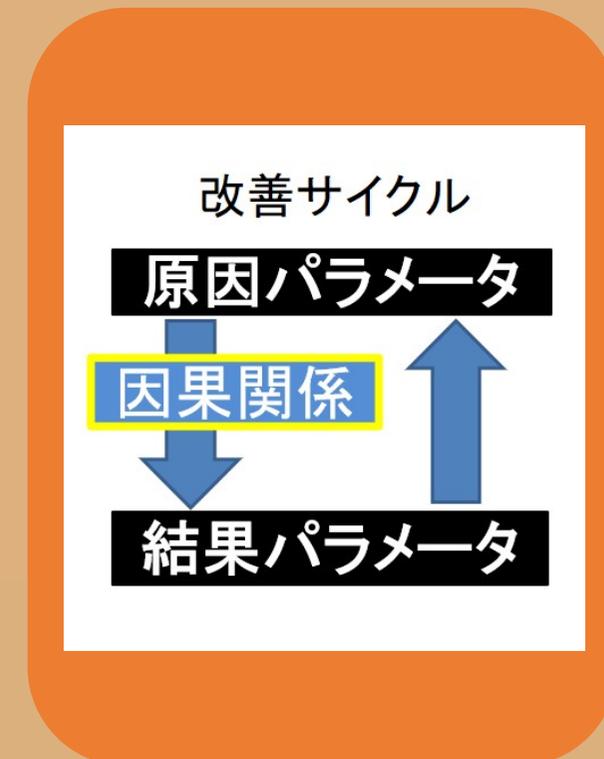
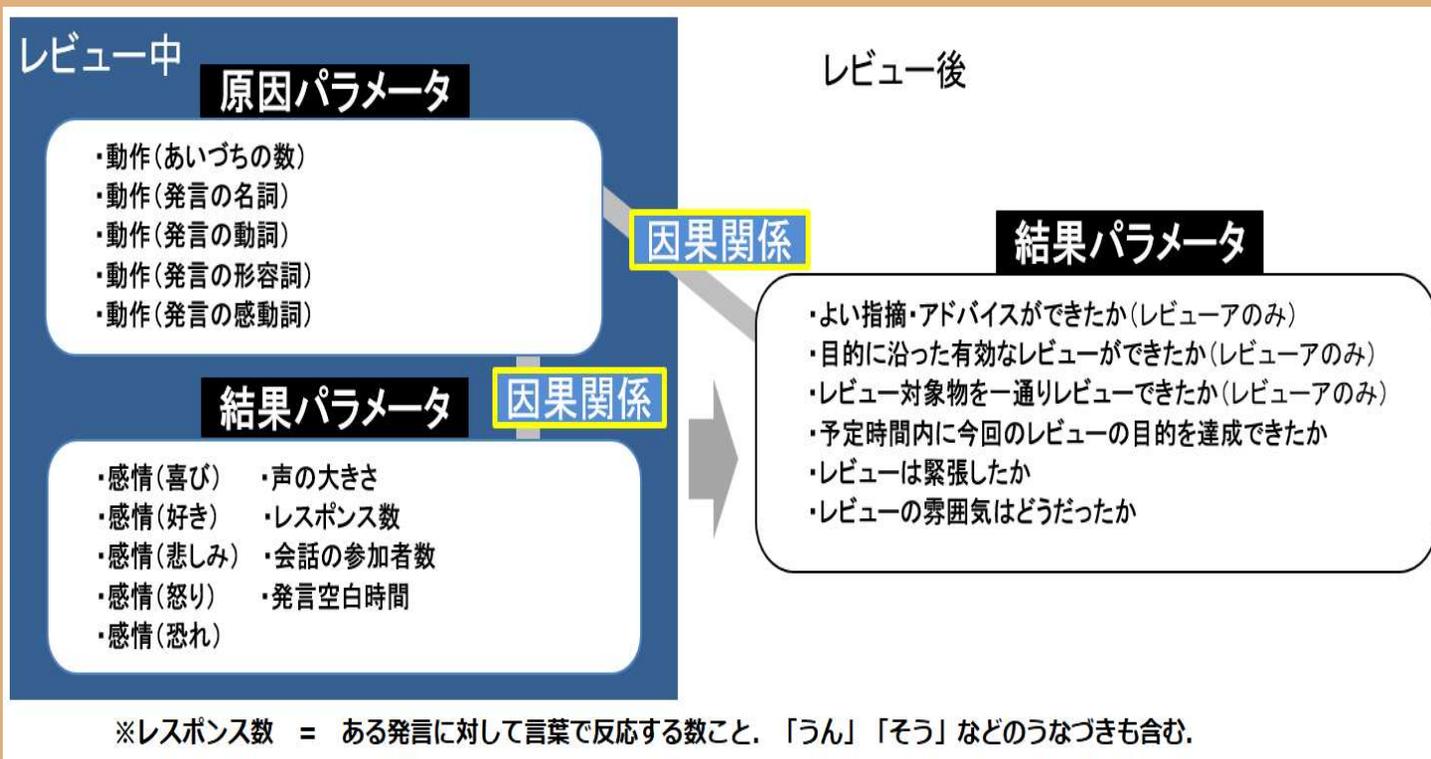
～レビュー会議の会話データを分析し、発言に関する心理的パラメータを測定～

「発言しやすさ」に影響を及ぼすいくつかのパラメータと  
それらのパラメータ間の関係性（因果関係）を探る!!!



なにを測りますか？

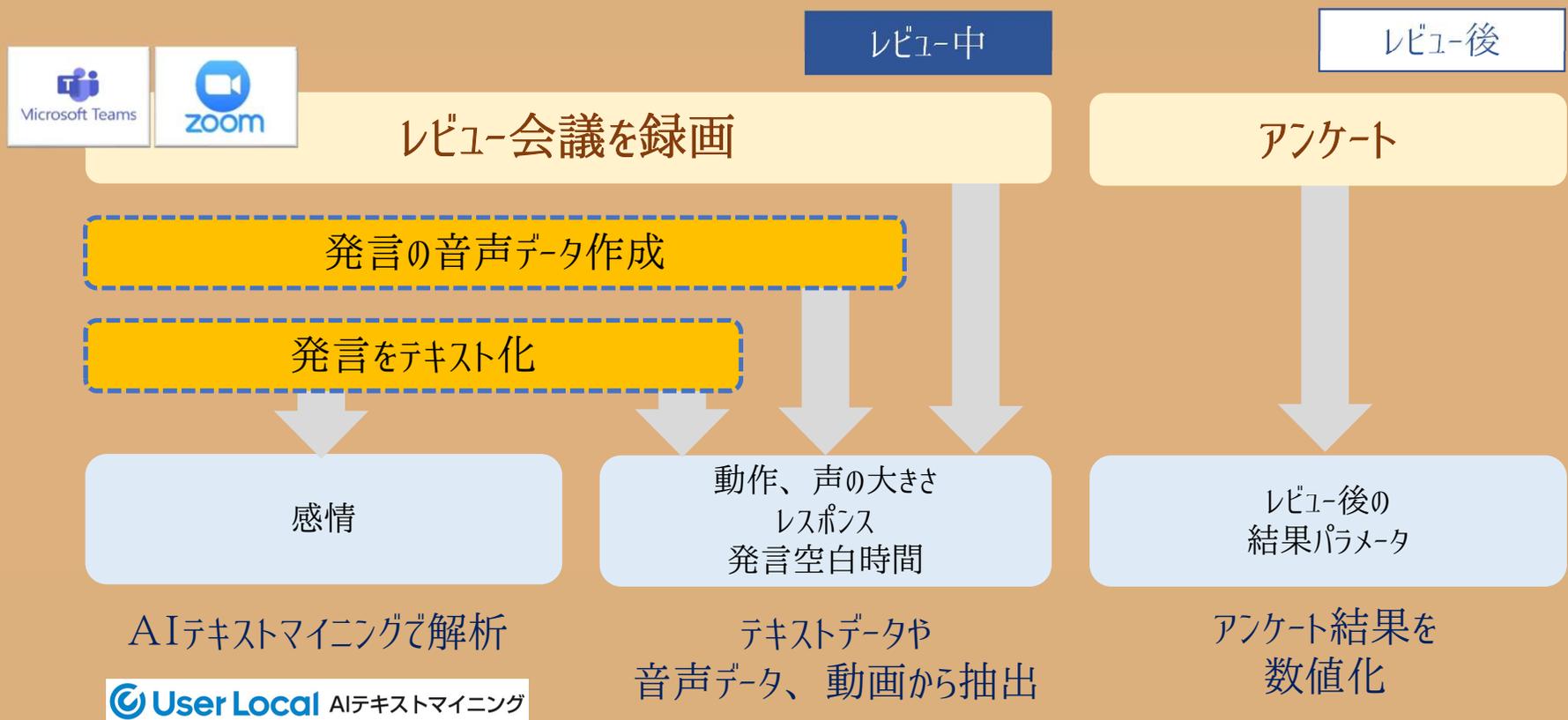
# エーオース (Ēōs) パラメータ





どうやって測りますか？

# パラメータ収集方法





# 実験の方法



# 予備実験

User Local AIテキストマイニング

- AIテキストマイニングによる測定の妥当性確認が目的
- 効果の低い／高いレビューを **意図的に作り出したシナリオ** を利用



測定メトリクス⇒

感情

測定結果⇒

ホワイトレビューの方が  
✓ 「恐れ」や「悲しみ」が少ない  
✓ 「好き」や「喜び」が多い

# 本番実験

- ロールプレイによる模擬レビュー
- 研究員が2つのチームに分かれて、レビューアのロールを担う
- それぞれのチームが

ホワイトレビュー または ブラックレビュー を意識して再現



すべてのエーオース (Eōs) パラメータ

測定メトリクス⇒

感情

動作、声の大きさ  
レスポンス  
発言空白時間

レビュー後の  
結果パラメータ



実験の結果



# 今回測定したパラメータと測定方法

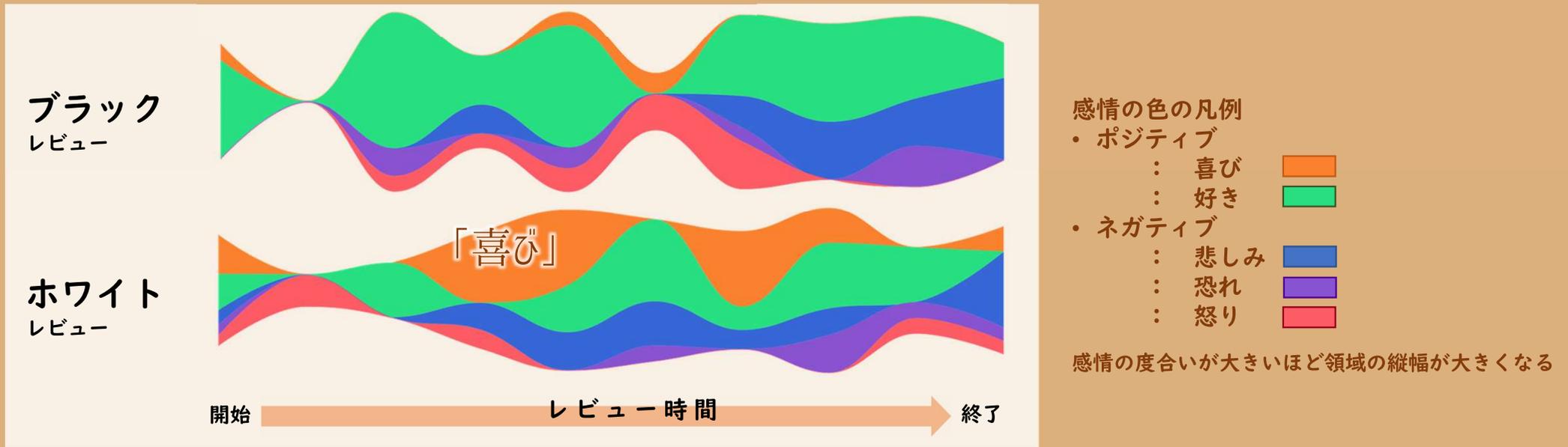
パラメータ		測りやすさ	測定方法
1	感情	◎(AI分析)	文字起こし結果をユーザーローカル社AIテキストマイニングツールで分析
2	あいづち	◎(テキストマイニング)	文字起こし結果をユーザーローカル社AIテキストマイニングツールで分類
3	発言に含まれる単語	◎(テキストマイニング)	
4	声の大きさ	△(音声編集ソフト)	録音結果を音声編集ソフトで発言ごとのピーク音量を測定
5	会話参加者数とレスポンス数	△～○(ツール性能に依る)	文字起こし結果およびそのタイムスタンプを元に集計(話題は音声から手分類)
6	発言空白時間	△～○(ツール性能に依る)	
7	参加者の満足度	○(Webアンケート)	参加者にWebアンケートを依頼し、アンケート結果をスコアリング・集計
	参加者の緊張度	○(Webアンケート)	
	参加者が感じた雰囲気	○(Webアンケート)	

測りやすさ◎はツール使用のみ、○は簡単な集計があるもの、△は一定の手作業が発生するもの

# 1 感情

～感情分析 powered by ～  
(<https://textmining.userlocal.jp/>)

レビュー中の **レビューイの感情推移** をAIツールを用いて 可視化



▶ 「喜び」の感情がホワイトレビューで大きく出現している

## 2 あいづち

レビュー中の発言から、参加者のあいづちの数 を計測 (回)

あいづち	ブラックレビュー		ホワイトレビュー	
	レビューイ	レビューア	レビューイ	レビューア
はい うん そう そうですね なるほど 確かに } を含む会話の件数	31	10	27	33

3.3倍

- ▶ 「レビューア」のあいづちの数 が ホワイトレビューはブラックレビューの3倍超

# 3 発言に含まれる単語

レビュー中の発言から **品詞の数** を計測

名詞 (上位5種を表示)

ブラックレビュー(11種35語)			ホワイトレビュー(12種42語)		
#	単語	出現数	#	単語	出現数
1	画面	7	1	申請	8
2	入力	6	2	ステータス	4
3	申請	5	3	承認	4
4	ステータス	3	4	確か	4
5	作成	3	5	画面	3

動詞 (上位5種を表示)

ブラックレビュー(15種37語)			ホワイトレビュー(15種37語)		
#	単語	出現数	#	単語	出現数
1	入れる	9	1	思う	12
2	できる	7	2	書く	4
3	書く	5	3	終わる	3
4	わかる	4	4	隠す	2
5	入る	2	5	わかる	2

形容詞 (上位3種を表示)

ブラックレビュー(3種6語)			ホワイトレビュー(8種15語)		
#	単語	出現数	#	単語	出現数
1	いい	3	1	いい	7
2	面倒くさい	2	2	すごい	2
3	新しい	1	3	嬉しい	1

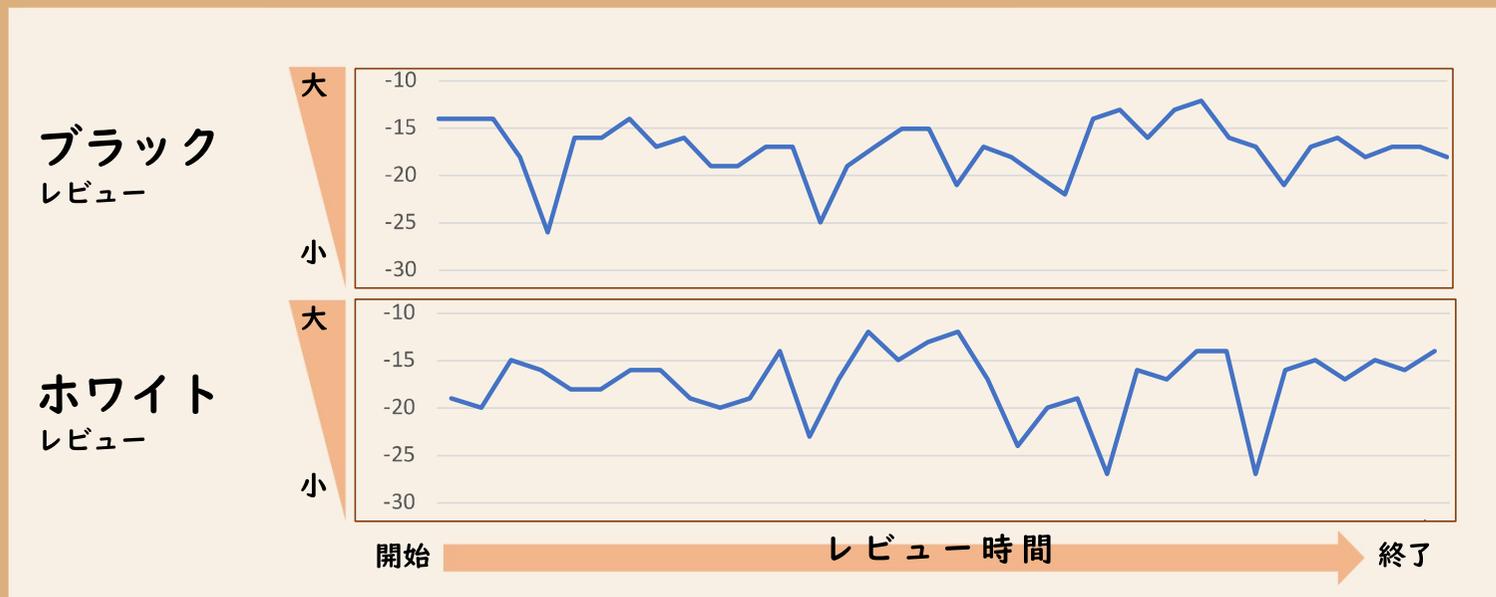
感動詞 (全数を表示)

ブラックレビュー(3種6語)			ホワイトレビュー(3種5語)		
#	単語	出現数	#	単語	出現数
1	なるほど	4	1	そうですね	3
2	そうですね	1	2	なるほど	1
3	ありがとう	1	3	ありがとう	1

▶ ホワイトレビューで動詞「**思う**」が突出。また形容詞では **ポジティブな単語** が多い

## 4 声の大きさ

レビュー中のレビューイの音量変化を計測



のるしかない  
このビッグウェーブに

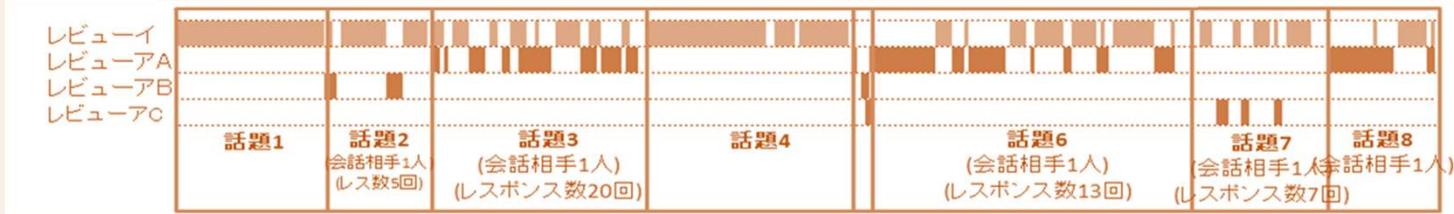


▶ 指摘タイミングで音量が下がる傾向は共通。その他で目立った差はみられず

# 5 会話参加者数とレスポンス数

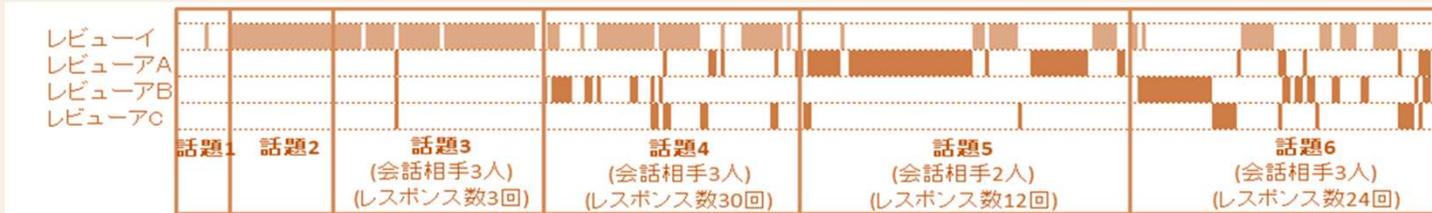
一つの話題ごとの 会話参加者数 とレスポンス数 を計測

ブラック  
レビュー



会話参加者数：1  
レスポンス数：~20

ホワイト  
レビュー



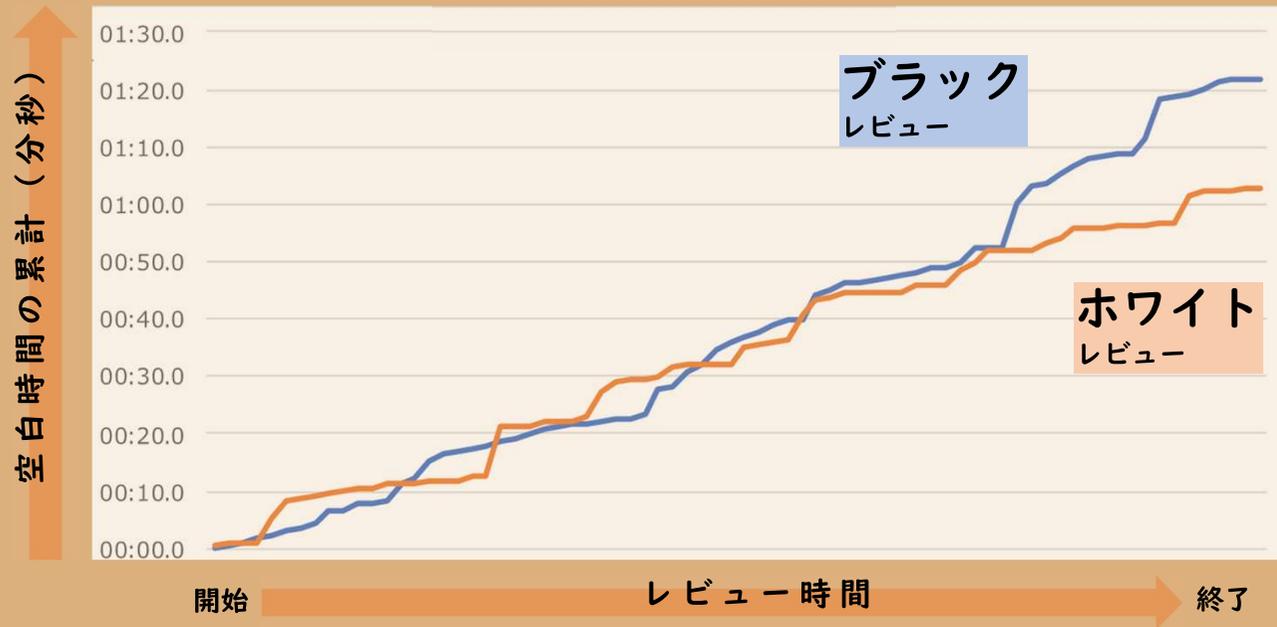
会話参加者数：3  
レスポンス数：~30

開始  レビュー時間 → 終了

▶ ホワイトレビューは「会話の参加者数」も「レスポンス数」も多い

# 6 発言空白時間

レビュー中の 発言間の 空白時間 を計測



▶ ホワイトレビューは「累計時間」が短い、「指摘後の空白時間」も短い傾向である

# 7 参加者アンケート

レビュー終了後、レビューアにアンケートを実施して **満足度** と **心理面** を確認

	アンケート項目		スコアリング				アンケート結果(合計)	
			3点	2点	1点	0点	ブラックレビュー	ホワイトレビュー
1	満足度 因子	よい指摘・アドバイスができたか	できた	どちらかといえばできた	どちらかといえばできなかった	できなかった	17	32
2		目的に沿った有効なレビューができたか						
3		レビュー対象物を一通りレビューできたか						
4		予定時間内にレビューの目的を達成できたか	できた	7割超達成できた	5割~7割達成できた	半分未満だった		
5	心理的 因子	レビューは緊張したか	—	緊張した	少し緊張した	緊張しなかった	3	0
6		レビューの雰囲気はどうだったか	—	厳粛な雰囲気	通常の会議の雰囲気	緩い雰囲気	5	0

各レビューに参加したレビューア3名のスコアを合計

▶ ホワイトレビューは「満足度」が高く、「緊張なし」「ゆるい雰囲気」と感じていた

# 実験結果のまとめ

パラメータ	有効性	実験で確認できた差異	測りやすさ
1 感情	◎	ホワイトレビューでレビューイの「喜び」の感情が大きい	◎(AI分析)
2 あいづち	◎	ホワイトレビューではレビューアのあいづち数がブラックレビューの3倍超	◎(テキストマイニング)
3 発言に含まれる単語	◎	ホワイトレビューで肯定的な形容詞が多い／「思う」等断定しない表現が多い	◎(テキストマイニング)
4 声の大きさ	要検証	今回はホワイトレビューとブラックレビューで差異点を確認できず	△(音声編集ソフト)
5 会話参加者数とレスポンス数	◎	ホワイトレビューがブラックレビューよりレスポンス数、参加者数ともに多くなった	△～○(ツール性能に依る)
6 発言空白時間	○	ホワイトレビューがブラックレビューより空白時間の合計が少なくなった	△～○(ツール性能に依る)
参加者の満足度	○	ホワイトレビューは満足度が高く、ブラックレビューは低い	△(Webアンケート)
参加者の緊張度	○	ホワイトレビューは緊張していない。ブラックレビューは参加者により異なる	△(Webアンケート)
参加者が感じた雰囲気	○	ホワイトレビューは緩い雰囲気、ブラックレビューは通常または厳粛な雰囲気	△(Webアンケート)

殆どのパラメータでブラックホワイトでの差異あり。特に有効性◎は「発言しやすさ」測定に有効

# 今後の展望

新たなレビュー品質メトリクス「Eōs」の扉は開かれたばかり!!!  
更に有効なパラメータや因果関係が発見される可能性を秘めている!!!

## ● データ収集・分析の継続

パラメータ間の新たな因果関係が見つかり、「発言しやすさ」のコントロール手段を増やすことができる

深  
める

広  
げる

## ● 新たなパラメータの追加

原因パラメータ：

- ・ ジェスチャー(身振り手振り, 身体の動き)
- ・ 顔の表情(笑顔, 顔色)など

結果パラメータ：

- ・ バイタルデータ(心拍数, 体温, 発汗など)
- ・ 声の波形(音程, 震え)など

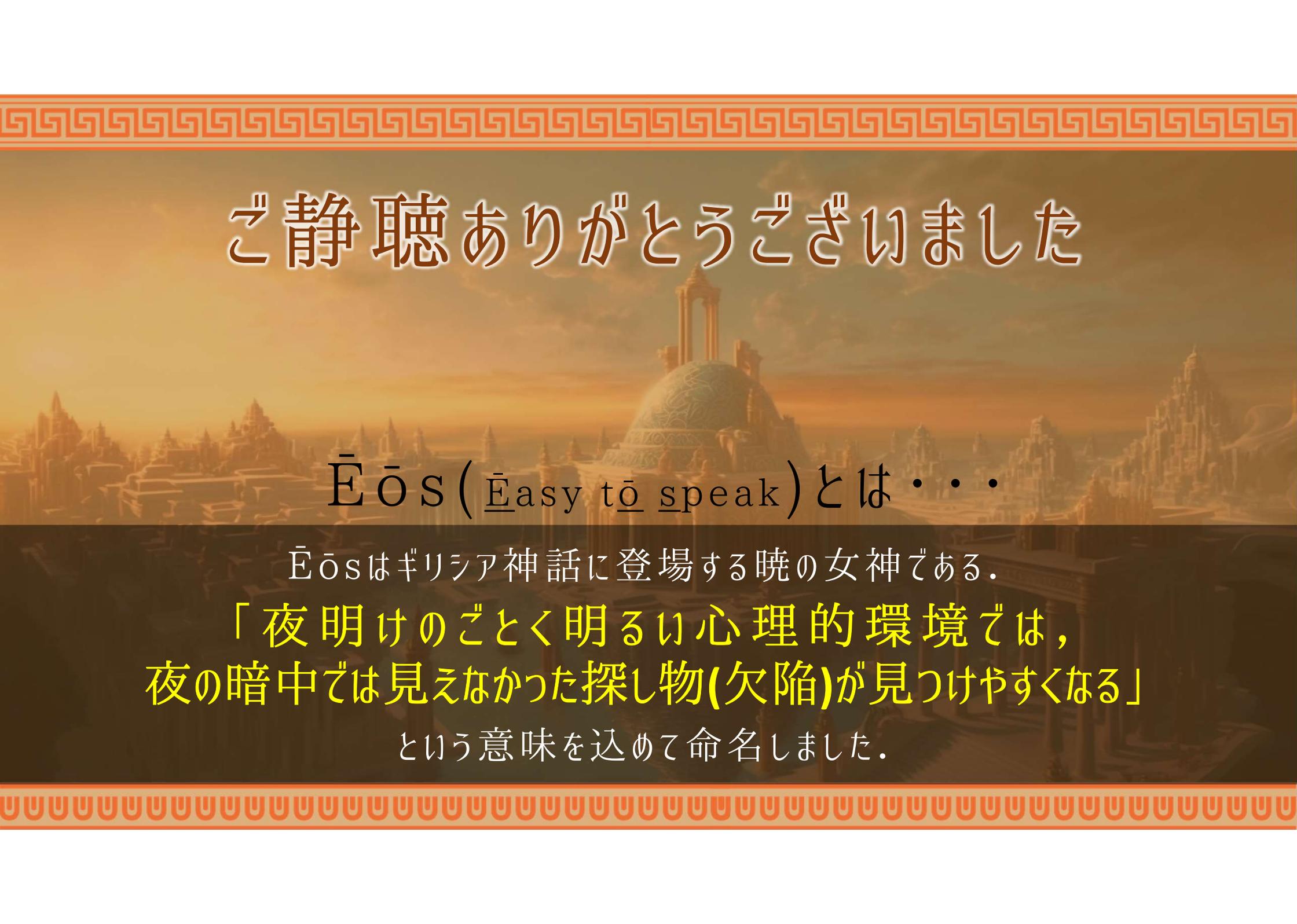


更なる研究が求められる!!!



この研究に出てくる「エーオース」って、





ご静聴ありがとうございました

Ēōs (Ēasy tō speak) とは・・・

Ēōsはギリシア神話に登場する暁の女神である。

「夜明けのごとく明るい心理的環境では、  
夜の暗くでは見えなかった探し物(欠陥)を見つけやすくなる」

という意味を込めて命名しました。