

Author's accepted manuscript of Liu, N. (2022). Speaking in the first-person singular or plural: A multifactorial, speech corpus-based analysis of institutional interpreters. *Interpreting*, 1–35. <http://doi.org/10.1075/intp.00088.liu>

Appendices available from <http://doi.org/10.1075/intp.00088.liu.additional> and data and code available from <https://osf.io/z5gtr/>.

Erratum: The citation of behavioural profile in section 4 should be Divjak, D. & Gries, S. Th. (2006). Ways of trying in Russian: clustering behavioral profiles. *Corpus Linguistics and Linguistic Theory* 2 (1), 23–60. <https://doi.org/10.1515/CLLT.2006.002>

## **Speaking in the first-person singular or plural: A multifactorial, speech corpus-based analysis of institutional interpreters**

Nannan Liu

University of Hong Kong

Author's address

Nannan Liu

School of Chinese

University of Hong Kong

Pokfulam

Hong Kong

nnl93@connect.hku.hk

### **Biographical note**

**Nannan Liu** received a PhD in Translation from the University of Hong Kong in 2020. Her research interests are interpreting studies, computational linguistics and cognitive linguistics. She is the author of the Python library *muldichinese*, a register analyser.

### **Acknowledgements**

I am grateful to Christophe Coupé for comments on an earlier draft of this article and to Pei-yin Lin for enlightening discussions. My profound gratitude goes to the editors and the three reviewers for their genuine and constructive suggestions for improving this article.

# Speaking in the first-person singular or plural: A multifactorial, speech corpus-based analysis of institutional interpreters

## **Abstract**

Conventionally, the professional interpreter speaks in the first-person singular. Research in corpus-based discourse analysis has reported shifts from this norm towards first-person plural becoming the most frequent pronoun shift in political institutional interpreting, possibly signalling interpreter ‘alignment with the institution’. Nonetheless, few studies have teased apart the simultaneous constraints of social, cognitive, and linguistic factors on institutional interpreters’ preference for the plural. The present research adopts the usage-based theory to consider the three types of explanations together. It extends recent multivariate methodologies based on this theory to analyse 2,438 first-person cases in parallel interpreting and comparable speech corpora. Following robust context analyses and cross-linguistic prosodic transcription, this study weighs the strengths of 33 associates regarding the three explanations through regression analyses. The results show that first-person shifts are better explained by chunking effects when interpreters process complex forms and referents in the source and target speeches, and when they process zero-subject source inputs. The institutional alignment explanation fails to account for the extensive grammaticalisation of plural constructions in the interpreted speech. When all the interactive and additive effects are considered together, institutional alignment or monofactorial paradigms have little explanatory power. This study concludes by highlighting the relevance of usage-based multifactorial designs to interpreting research.

**Keywords:** first-person interpreting, speech corpus, usage-based theory, grammaticalisation, MuPDAR(F)

## **1. Introduction**

The norm that the interpreter speaks “in the first person as if s/he was the orator” (Harris 1990: 115) expects professional interpreters to speak in the first-person singular. In interpreting for political institutions such as the European Parliament (EP) and the Chinese premier’s press conferences (CPPCs), departures from this norm to first-person plural or the introduction of such a plural that is ‘not triggered’ by source speech have been reported as the most frequent

category of pronoun shift (Beaton-Thome 2010; Gu & Tipton 2020; Li & Wang 2012), as shown in Example 1 (first-person forms underlined). Such shifts may be a sign of the interpreters' alignment with the institution and they have implications for the way ideology is expressed in international politics.

### Example 1.

Original Chinese: 中美都表示 / 要维护 / 地区的 / 稳定 / 促进繁荣 // 我希望啊 / 言行一致 //

Gloss: (Both China and the United States expressed a commitment to regional stability and prosperity. I hope such talk will be matched by actions.)

Interpreted English: Both countries have stated / that we want to uphold / regional stability / and promote regional prosperity // We hope / that these commitments / will be reflected / in action //<sup>1</sup>

Research into shifts towards plural first-person pronouns (FPPs) has generally assumed that such shifts are optional and they depend on whether the speaker or interpreter intends to strengthen their personal obligation or identification with the political institution. In other words, there is a choice between “how [speakers/]interpreters position themselves as individuals” (i.e. ‘I’) or as “a group within the institution” (i.e. ‘we’; Beaton-Thome 2010: 119). FPP shifts have been explained in terms of institutional alignment (Beaton-Thome 2010; Gu & Tipton 2020), cognitive effort (Li & Halverson 2020; Plevoets & Defrancq 2018) and cross-linguistic influence (Li & Wang 2012). Most studies have relied on ad hoc examples, non-statistical designs and source–target comparisons, which have conceptual and methodological limitations.

Conceptually, such designs are often monofactorial in that they tend to over-analyse the effects of one predictor (institutional background) on the response (FPP use), while sidelining the additive and interactive impact of other predictors. I reiterate here De Sutter and Lefer's (2020: 1) call for a multifactorial research agenda in that translation and interpreting are “an inherently multidimensional linguistic activity and product, which is simultaneously

---

<sup>1</sup> All the examples in this article are taken from the CPPC interpreting and comparable corpora used in the present study (see section 4.1).

constrained by sociocultural, technological and cognitive factors”. Disregarding the myriad factors behind an interpreter’s choices leads to a misconstrual of the effects of interpreting and to generalisations that are only slightly better than random guessing, as will be shown in section 5.1.

Methodologically, the extant research largely under-uses the affordances of corpus data. In this article, I focus on three aspects of this under-exploration: context, multifactoriality, and orality. First, a conference is a holistic communicative event, but most studies of FPP shifts aggregate the “text, situation, culture, and the entire course of action in a professional interpreting assignment” to frequency tallies, while ignoring the interpreted speech’s intracultural fit with recipient conventions (Pöchhacker 1995: 33–38). Translated into operationalisations, the studies often do not consider that (i) a native speaker of the recipient language may use a similar number of plural FPPs, if not more, in a context comparable to that faced by the interpreter, and (ii) the context described in the interpreted speech may favour the use of plural over singular. Therefore, both the intra- and the intercultural contexts need to be accounted for prior to ideological motivations being discussed.

Second, an empirically justified model should consider the ways in which a single factor is “enhanced, mitigated, maybe even offset by the presence of another” (Toury 2004: 25). However, the absence of advanced statistics in previous research precludes an assessment of the strength of alignment against all other predictors taken together.

Third, the use of prosody can “increase the ideological distancing (or alignment) on the part of the interpreter” (Beaton-Thome 2013: 387), but previous research has relied on transcripts rather than their sources, namely audiovisual recordings or on-site observation. Granted, prosodic transcription remains a quandary in corpus-based studies, particularly because of the lack of agreed-upon conventions. Advances in automatic transcription (used in Lenglet & Michaux 2020) and cross-linguistic frameworks can help to fill this lacuna, but they have not been used in research into pronoun shifts.

This article presents a speech corpus-based multifactorial analysis of those factors that condition institutional interpreters’ choices between first-person singular and plural pronouns. Conceived from the maximalist perspective of usage-based and grammaticalisation theories, this research seeks to re-create the context of political conferences beyond mere frequency tallies, on the assumption that first-person shifts emerge from interpreters’ language use in such contexts (Bybee 2010; Langacker 1988). I extend recent multivariate methods – namely,

Behavioural Profile (BP) and Multifactorial Prediction and Deviation Analysis using Regression/Random Forests (MuPDAR[F]; Gries 2010; Gries & Deshors 2014) – to consider social, cognitive, and linguistic explanations for FPP shifts simultaneously. The study pursued four goals:

1. an exploratory one, namely, identifying the factors that influence the choices of interpreters in political institutional settings between first-person singular and plural;
2. a confirmatory one, namely, disentangling the main explanations proposed for the FPP plural preference of institutional interpreters;
3. a methodological one, namely, illustrating one way of using multivariate methods to characterise the internal and external contexts of interpreters' choices;
4. a theoretical one, namely, exemplifying the relevance of the usage-based theory to an understanding of interpreting as a multidimensional communicative process and product.

## **2. Explanations of institutional interpreters' FPP plural preference**

Three main explanations can be deduced from previous research into institutional interpreters' FPP use: social, cognitive, and linguistic. Each of these is described below.

### **2.1 *Social explanations***

Beaton-Thome (2010: 135) proposed a plural preference among EP interpreters in contrast to that of source speakers as a result of the mediators' "ideological positioning as part of the institution". The positioning factor was alluded to by Monacelli (2009) and developed in Beaton-Thome (2013), Fu and Chen (2019), and Gu and Tipton (2020). Based on close readings of corpus transcripts, they suggested a trend in institutional interpreting of shifts from source singular FPPs, noun phrases (NPs), zero FPPs, passive syntax, and 'they' to plural FPPs.

The exact referent of such plurals was "unclear from the context", but it was understood as indexing the dominant (supra)national institutions that interpreters work for (Beaton-Thome 2010: 134; Fu & Chen 2019; Gu & Tipton 2020). It was argued that such shifts strengthened the institutional presence while marginalising speaker-personal or regional voices (Beaton-Thome 2010; Gu & Tipton 2020). Theories of discourse meaning were invoked to pinpoint the plural's 'ideological salience' in presupposing and reinforcing "in-group membership" or an

“us versus them” differentiation (van Dijk 2004: 730). On this view, shifts towards the plural were prompted by (i) topics that encode such ideological salience – for example, the Israel–Palestine relationship (Monacelli 2009) or the Guantánamo Bay detainees (Beaton-Thome 2010, 2013) – and (ii) institutional loyalty: references to the institution or to authoritative positions were “replicated” by staff interpreters, whereas references to non-affiliated identities were “weakened” or omitted (Beaton-Thome 2010: 135; Fu & Chen 2019: 28).

Cognitive, (para)linguistic, and situational factors were only cursorily mentioned in these discussions. For instance, Beaton-Thome (2010, 2013) suggested that ‘we’ was used to reduce processing effort in simultaneous interpreting (SI). Based on post hoc analyses she argued that prosodic stress, hesitation, filled pauses, and self-corrections revealed the ideological positioning of interpreters (Beaton-Thome 2013) and that the “situational settings of SI” might drive the choice of *wir alle* (‘we all’; Beaton-Thome 2010: 127). Overall, though, this line of reasoning sidesteps non-ideological factors to the extent of claiming that “none of the additions [of plural FPPs] were triggered by the ST [source text] or the grammatical differences between the two languages” (Gu & Tipton 2020: 414).

## 2.2 *Cognitive processing*

The cognitive demands posed by split-second processing and production affect an interpreter’s preferences, including the surface forms and the underlying referents of FPPs. Regarding forms, research on formulaic sequences has identified plural constructions as the most frequent sequences in institutional settings (Li & Halverson 2020; Plevoets & Defrancq 2018). Using frequency-based criteria, Plevoets and Defrancq (2018) identified *il nous faut* (‘we need to’) as a formulaic sequence in EP French source speech. Sequences in the source and target are easy to process and produce, leading to a low rate of disfluencies in interpreted speech (Plevoets & Defrancq 2018). Canvassing a range of causal factors in formulaicity, Li and Halverson (2020) demonstrated the association between priming and CPPC interpreters’ selection of FPP sequences. Priming refers to speakers’ tendency to re-use the structures that they have recently encountered or produced, such repetition facilitating processing. According to Li and Halverson (2020), plural sequences can derive from self-priming (production–production) or other-priming (source comprehension–[note-taking]–target production). For instance, other-priming from the vague plural lemma that is prevalent in Cantonese was suggested to have caused the

preference of Hong Kong government interpreters for using the plural in SI into English (Li & Wang 2012).

A methodological limitation of previous formulaicity research is that the sequences were extracted from transcripts based on “somewhat arbitrary” frequency cut-offs (Li & Halverson 2020: 11). A more robust method lies in the prosody–discourse interface. Based on analyses of the International Corpus of English, Dehé and Wichmann (2010) show that the singular FPP and epistemic verb construction (e.g. ‘I think’, ‘I believe’) has multiple functions on a continuum from the propositional to the formulaic, which are reflected in their prosodic realisation. In specific terms, when the singular FPP bears an accent (see section 4.2.4), the construction expresses the speaker’s true belief, whereas when the adjacent verb receives prominence, it provides a comment on the rest of the sentence. In cases where the construction joins neighbouring intonational phrase(s) and receives no prominence, it has grammaticalised into a formula with no semantic value (Dehé & Wichmann 2010). This inverse relationship between prosodic prominence and grammaticalisation is compatible with theories of intonational meaning which posit a correlation between semantic weight and prosodic prominence (Pierrehumbert & Hirschberg 1990). By Beaton-Thome’s (2013) account, an ideologically salient form is delivered with stress or disfluencies and is therefore unlikely to have grammaticalised.

As regards underlying referents, shifts towards plurals indexing dominant institutions hark back to the bondedness hierarchy, which describes modifier–noun relations (Croft 2003; cf. Beaton-Thome 2010: 122 on “hierarchical” referents). This hierarchy is represented as “article > deictic > interrogative > quantifier > adjective > relative clause”, with the bondedness to the modified noun decreasing from the article to the relative clause (Croft 2003). Hawkins (2004) predicts greater advantages in online NP processing with definite-article rather than adjective modifiers, owing to the definite article’s high degree of morphosyntactic bondedness to the noun and grammaticalisation. For instance, relative to the source NP “the Czech Presidency” (Beaton-Thome 2010: 131), which is modified by an adjective, the interpreter’s choice of “the Council” underlying the plural is modified by an article and associated with the highest degree of bondedness and processing ease.

### **2.3 *Linguistic diversity***

Languages such as Mandarin and Spanish allow zero subjects, and so when interpreting into subject-obligatory languages, shifts such as those towards FPPs are necessary. Li and Halverson (2020) show that in cases of subjectless clauses in source speeches, CPPC interpreters favour the plural formulaic sequences ‘we will continue to’ and ‘we are going to’ over the singular and the plural lemmas. The linguistic explanation therefore aligns with cognitive explanations.

In a nutshell, the choice between the singular or the plural FPP in institutional interpreting exemplifies the simultaneous constraints of social, cognitive, and linguistic factors on the interpreting product. As insightfully pointed out by Li and Halverson (2020: 4), it is incumbent on empirical research to “ultimately tease the potential factors apart and test their relative strengths and weaknesses”, with the selection of units of analysis motivated by a specific theoretical framework. In the present study, this goal is pursued through a multifactorial, usage-based approach.

### **3. A usage-based approach**

The present research adopts the usage-based approach because it aims to consider the three explanations together. Following Croft (2003), I consider corpus-internal explanations (e.g. the alignment of institutional interpreters and formulaic sequences) inferior to external explanations of general human characteristics. The basic tenet of usage-based theory is that language structure emerges from language use: the repeated use of the same sounds, words, and patterns has an impact on the cognitive storage and processing of linguistic experiences that give a language its structure, or grammar (Bybee 2010; Langacker 1988).

Grammar derives from domain-general cognitive processes such as categorisation, chunking, and inferencing, which are common to all languages and operate whenever people are speaking (Bybee 2010). Of relevance here is chunking, or the sequential relationships fostered when two or more words are often produced together (Bybee 2010). A few repetitions of word combinations result in a weak chunk, whereas more frequent chunks are easily accessible as wholes in cognition (Langacker 1988). Formulaic sequences, which result from chunking, contribute significantly to fluency and ease (Bybee 2010).

Chunking effects are evidenced by an increase in the frequency of co-occurrence, an association with repetition or priming, production fluency, component meaning loss, and



morphosyntactic bondedness. Chunking is also associated with the loss of phonetic segments, the ability to receive stress, and phonological independence (Bybee 2010).

Chunking is among the cognitive mechanisms that propel grammaticalisation (Bybee 2010; Narrog & Heine 2021). Grammaticalisation refers to the process by which lexical forms acquire grammatical functions or by which grammatical forms develop new grammatical functions “in specific contexts and constructions” (Narrog & Heine 2021: 1). Because of the underlying role of chunking in grammaticalisation, the same diagnostics outlined above for chunking also apply to grammaticalisation (Bybee 2010; Narrog & Heine 2021).

Usage-based and grammaticalisation theories provide a comprehensive framework with which to analyse the context of FPP choices. In the light of the literature and the theoretical models, the present study examines the following hypotheses:

1. By the institutional alignment account,
  - a) the first-person plural would increase in frequency, attract stress or disfluencies, and display a lower degree of grammaticalisation vis-à-vis the singular when speakers or interpreters signal alignment with institutional positions on ideologically salient topics across original Chinese (OC), interpreted English (IE), and native English (NE);
  - b) the plural in IE would be less grammaticalised than that in OC, because interpreters had strengthened institutional alignment on ideologically salient topics compared to source speakers;
2. By the cognitive processing explanations,
  - a) plural constructions would be more grammaticalised than the singular across the three varieties when speakers process and produce complex forms and referents and interpreters do so in the source and target;
  - b) plural constructions in IE would be more grammaticalised than those in OC and NE owing to the greater cognitive demands of interpreting;
3. By the linguistic diversity account, plural constructions would be more grammaticalised than the singular in IE when there is no subject form in the source segment, because interpreters may have developed a plural formula for dealing with zero-subject inputs.

## 4. Data and methods<sup>2</sup>

This study extends the BP and MuPDAR(F) approaches developed by Gries and colleagues (Gries 2010; Gries & Deshors 2014)<sup>3</sup> to achieve its exploratory, confirmatory, and methodological goals. Utterances containing FPPs ('FPP cases') were extracted from speech corpora, annotated for an array of contextual variables (section 4.2) and then analysed statistically (section 4.3).

### 4.1 *Speech corpora*

FPP cases were extracted from parallel CPPC interpreting and comparable speech corpora (Liu 2020). The CPPCs were held after sessions of the National People's Congress and Chinese People's Political Consultative Conference. They were moderated by Congress spokespersons and attended by journalists, who asked the premiers questions. FPP cases which have maximal audio quality were sampled in original Chinese (OC) and interpreted English (IE) in the events of 2004–2006 and 2013–2015. OC data include moderator housekeeping notes, reporter questions in Mandarin, and premier statements on issues related to mainland China, international relations, and Taiwan. The IE data are consecutive interpretations made by five different staff interpreters of the Chinese Ministry of Foreign Affairs, who worked into their B language and were anonymised as IE1, IE2, IE3, IE4, and IE5.

Liu (in press) finds the register of former US President George W. Bush's State of the Union (SOTU) addresses to be the most similar to that of IE compared to 16 other registers. For this reason, comparable native English (NE) cases were obtained from Bush's SOTU addresses from 2003 to 2008, covering US issues, the Iraq War, and international relations apart from that war. Audio recordings of the SOTUs were segmented into files with applause removed. The selection of the addresses also minimises the differences in the delivery mode, the time frame (from 2003 to 2015), and the speaking time of the primary speakers between the parallel and comparable corpora. Video observation and previous research (Liu, in press) show that at the CPPCs, premiers, reporters, and moderators spoke from prepared scripts or with the aid of notes. As the SOTU addresses were read out from the teleprompter and the IE was

---

<sup>2</sup> The annotated dataset and analysis code are available at <https://osf.io/z5gtr/>.

<sup>3</sup> I am deeply grateful to Stefan Th. Gries for the Behavioural Profile for R script and explanations of MuPDAR(F) procedures.

reconstructed from notes, the utterances were read out or delivered with the aid of notes across the three varieties. The premiers and the US President spoke for about 50–60 minutes per session.<sup>4</sup> Orthographic transcripts were aligned with the corresponding audio files using the Penn Phonetics Lab Forced Aligners (Yuan & Liberman 2008).

Because of the absence of any statistical testing in research that reports the plural preference (section 2.1), the analysed samples from the corpora were first used to examine the FPP distribution as a function of VARIETY (i.e. OC, IE, and NE). Next, validation samples comprising OC and IE utterances in 2003, 2007, 2016, and 2017, along with the SOTU addresses by Barack Obama in 2013–2016, were selected from the same corpora (Liu 2020). The effect of VARIETY on the FPP distribution in the samples was analysed. This step aimed to examine whether the plural preference is sensitive to the parts of the corpora used. The two steps constitute a ‘mini-replication’ in this study. The analyses proceed if the plural preference is robust. An overview of the corpora is provided in Table 1.

**Table 1.** Corpora composition

Corpus	Variety	Sample type	<i>N</i> of tokens	Audio length
Parallel	OC	Analysis	48,926	5 h 36 min 7 s
		Validation	37,403	
	IE	Analysis	38,820	4 h 33 min 37 s
		Validation	27,237	
Comparable	NE	Analysis	32,520	3 h 40 min 36 s
		Validation	27,204	
Total			212,110	13 h 50 min 20 s

The prosodic position of a form influences its phonetic realisation: across many languages, phrase-initial syllables are lengthened and strengthened whereas phrase-medial syllables are shortened and reduced (Cole 2015). To eliminate the confounding role of position in phonetic erosion, the analyses considered phrase-initial FPP subjects. Multifactorial studies assume that all cases are freely variable (Kruger & De Sutter 2018) between the singular and the plural, and

---

<sup>4</sup> I am not aware of political presentations in English varieties other than American English which show such similarities with CPPCs.

cases that impose an FPP form (i.e. only one of the forms is acceptable; for instance, the first FPP in the IE of Example 1 refers to “both countries” and has to be “we”) or involve non-subject, filler, or archaic usage were discarded. The dataset comprises 2,438 instances of FPPs (OC: 637, IE: 1,051, NE: 750).

#### 4.2 Behavioural Profile (BP) annotation

BP is used to assemble a list of near-exhaustive variables to describe the language-internal and -external context of FPP choices. Rather than applying exploratory methods following the annotation (cf. Gries 2010), the dataset is fed into regression analyses to achieve the confirmatory goal of the present study. BP is compatible with the tenet of usage-based theory and the proposal that chunking can affect all levels of grammar (Bybee 2010; Gries 2010). For the purposes of this study, it is considered an attempt at “holistic documentation” (Pöchhacker 1995: 41). All 2,438 cases were coded for the 33 variables listed in Table 2.

**Table 2.** Annotated variables and their relationship with the hypotheses

Hypothesis	Number	Type	Variable	Levels
1, 2, and 3	1	Morphological	Form	Singular, plural
	2	Variety	Variety	OC, IE, NE
	3	Random-effects	Audio ID	OC: 317, IE: 309, NE: 408
	4	Random-effects	Interpreter	IE1, IE2, IE3, IE4, IE5
	5	Random-effects	Year	OC and IE: 2004, 2005, 2006, 2013, 2014, 2015
				NE: 2003, 2004, 2005, 2006, 2007, 2008
	6	Random-effects	Lemma	OC: 184, IE: 267, NE: 225
7	Random-effects	Match	OC: 313, IE: 509, NE: 409	
1	8	Social	Topic	OC and IE: mainland China (CN), international relations (INT), Taiwan (TW), procedural housekeeping (PH), questions (Q) NE: US, Iraq War (IW), INT
2	9	Mediation-related	Mediation status	Non-interpreting, interpreting
	10	Morphological	Aspect	Simple, perfect, progressive
	11	Morphological	Tense	Present, past, future
	12	Morphosyntactic	Mood	Realis, irrealis
	13	Prosodic	Delivery rate	Number of syllables per second per audio ID
	14	Referential-morphosyntactic	Bondedness	Article/bare, deictic, adjective/quantifier, non-NP
	15	Referential-semantic	Animacy	Common noun, proper noun, person
	16	Referential-semantic	Concreteness	Abstract, concrete
17	Referential-semantic	Definiteness	Definite, indefinite	

	18	Referential-semantic	Group	Inclusive, exclusive
	19	Referential-semantic	Singularity	Plural, singular
	20	Semantic	Negation	Affirmative, negative
	21	Semantic	Modality	Unmarked, prediction, obligation, volition
	22	Semantic	Verb type	Activity, aspect, communication, cognition, emotion, existence, facilitation, wish
	23	Structural	Self-priming	No, yes
	24	Syntactic	Clause type	Main, coordinate, subordinate
	25	Syntactic	Sentence type	Declarative, non-declarative
	26	Syntactic	Voice	Active, passive
2 and 3	27	Structural	Other-priming	Other pronoun, NP, zero, plural FPP, singular FPP
1, 2, and 3	28	Prosodic	FPP prominence	Non-prominent, nuclear, prenuclear
	29	Prosodic	RPW1 prominence	Non-prominent, nuclear, prenuclear
	30	Prosodic	RPW2 prominence	Non-prominent, nuclear, prenuclear
	31	Phonetic and phonological	FPP break	0, 1, 2, 3, 4
	32	Phonetic and phonological	RPW1 break	0, 1, 2, 3, 4
	33	Phonetic and phonological	RPW2 break	0, 1, 2, 3, 4

The annotated dataset was explored to avoid (quasi-)complete separation, collinearity, and sparsity that would render regression results unreliable. Complete separation is caused by variables with some values that perfectly predict the response; it was dealt with by removing the variables or conflating their levels (Levshina 2015). The continuous predictor DELIVERY RATE was logarithmically transformed following normality tests. Referent ANIMACY, CONCRETENESS, DEFINITENESS, and SINGULARITY were discarded because of a high degree of collinearity; their impact is represented by BONDEDNESS and GROUP. INTERPRETER and YEAR contributed little to the models owing to their sparsity and they were therefore excluded.

#### 4.2.1 Random-effects terms

Random variables comprise (i) the MATCH of cases (i.e. the predicate, such as *qiānshūle* [have signed], “are going to adopt”) nested within the levels of LEMMA (i.e. *qiānshǔ* [sign], “adopt”), which allows varying intercepts for both the lemma and the match to accommodate lexical biases, and (ii) varying intercepts for AUDIO ID to control for file-specific idiosyncrasies. With nested random-effects, regression models take into account the fact that each match occurs with only one lemma, which helps to satisfy the assumption of observation independence (Gries 2021). Lexical bias refers to the (dis)association by which a lemma or a match favours or avoids

an FPP form. It reflects the FPP's immediate lexical context and co-occurrence with the adjacent match or lemma.

#### 4.2.2 Social explanations-related variable

The variable TOPIC indexes institutional alignment or response to situational settings (in cases of housekeeping). The topics of Taiwan, the Iraq War, and international relations are considered ideologically salient and may favour the plural. Owing to the PRC government's assertion that Taiwan is an inalienable part of one China, the utterances of premiers on Taiwan were structured in the vein of in-group membership (e.g. "we firmly oppose Taiwan independence"). In contrast, Bush constantly invoked the 'us versus them' paradigm to legitimise his war on terror (e.g. "we must act before the dangers are upon us"). This binary could also be found in international relations (e.g. "we hope they [the Japanese government] will refrain ...").

Four orthogonal contrasts were coded for the parallel data to tease the social explanations apart, pitting utterances

- (i) by premiers about Taiwan against those about mainland China and international relations;
- (ii) by premiers about international relations against those about Taiwan and mainland China;
- (iii) by premiers against those by reporters; and
- (iv) by moderators against those by others.

For NE, utterances about the Iraq War were pitted against those about the US and international issues, and those about international relations against the rest. As in Gries (2021), the contrasts were scaled so that coefficients in the summary regression tables reflect differences in logits directly (see the tables in Appendices C, D, and E).

#### 4.2.3 Cognitive and linguistic variables

Variables 9–27 include SELF- and OTHER-PRIMING, and those assumed to index the complexity of the grammatical context (e.g. ASPECT and CLAUSE TYPE) and underlying referents (i.e. BONDEDNESS and GROUP). The term 'bare' in BONDEDNESS refers to definite NPs without determiners, because articles are not permitted in Mandarin. For GROUP, exclusive referents (e.g. "the government") are generally more bonded than inclusive ones (e.g. "our nation", which is modified by the deictic).

OTHER-PRIMING relates both to cognitive processing (in cases of the FPP, the NP, and other pronouns) and linguistic diversity (in cases of zero). Non-FPP pronouns such as ‘they’ are infrequent in the dataset and it is assumed that they take some effort to process. I define the corresponding source subject of an IE case as ‘zero’ if there is no subject form in its host intonational phrase (see section 4.2.4), assuming that the unit “expresses a single focus of consciousness” and reflects a processing unit (see Lin 2018: 49 and references there). Biber et al. (2021) and Chao (1968) served as the reference for coding VERB TYPE.

#### 4.2.4 Prosodic variables

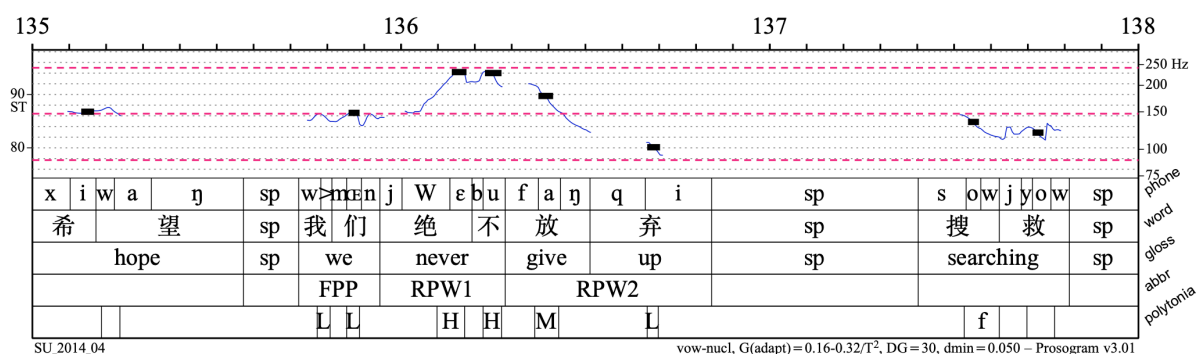
Variables 28–33 differentiate the degree of phonetic erosion in FPP constructions. The analyses examined the prominence types (related to stress loss) and the break indices (related to segmental loss and phonological integration) of FPP constructions, which comprise the FPP and its two right-hand prosodic words (PWs; labelled RPW1 and RPW2). An example of the FPP construction is provided in Figures 1 and 2. The Tones and Break Indices (ToBI) conventions were followed (Beckman & Elam 1997; Peng et al. 2005).

Prominence types were annotated based on the tone and stress systems of Mandarin and English. Almost every syllable in Mandarin has a tone specification and thus a local pitch contour. Therefore this language uses pitch-range expansion, duration lengthening, a fully realised lexical tone (denoted S3 in Mandarin ToBI; see the stress tier in Figure 2), and pitch-range compression of the following syllable to indicate prominence (Peng et al. 2005; Yang 2016). In English, the pitch accent marks syllable prominence, namely a high local pitch peak associated with a full unreduced vowel (Pierrehumbert & Hirschberg 1990). In both Mandarin and English, the rightmost accent in a prosodic phrase (see next paragraph) is nuclear and primary, whereas all accented syllables before the nucleus are prenuclear and secondary (Pierrehumbert & Hirschberg 1990). ToBI systems assign the diacritic \* to the pitch accent or the stressed syllable on the tone tier (see Figure 2), whether in nuclear or in prenuclear positions.

Break indices were coded by labelling the perceived level of disjuncture between the FPP and RPW1 (‘FPP BREAK’), RPW1 and RPW2 (‘RPW1 BREAK’), and RPW2 and the next prosodic word (PW) (‘RPW2 BREAK’) on an ordinal scale from 0 to 4 (see the break tier in Figure 2). A prosodic word (PW) such as “we’ll” and the Mandarin plural FPP *wōmen* contains a single word and associated unstressed functional words (Yang 2016). The break index numbers denote the following (Beckman & Elam 1997; Peng et al. 2005):

- The lowest level, 0, applies to cases of merged syllable boundaries, signifying the loss of phonetic segments and phonological independence (Narrog & Heine 2021).
- Break index 1 denotes a typical level of association in a fluent sequence.
- Break index 2 marks a pause or a pause-like separation.
- A prosodic phrase encompasses one or more pitch accents or stressed syllables and is marked by pre-boundary lengthening, pitch reset across phrases, a perceived major break, and a minor pause; it is assigned an index of 3 (Yang 2016).
- The host constituent of a prosodic phrase is an intonational phrase, which ends with pre-boundary lengthening, an audible pause, and a clear pitch reset; it is labelled 4.

Synchronised audio-transcription files were first analysed with Prosogram (Mertens 2014), with its Polytonia function applied to auto-transcribing the pitch level and the movement of each syllable. Figure 1 illustrates the results of the semi-automatic analysis.



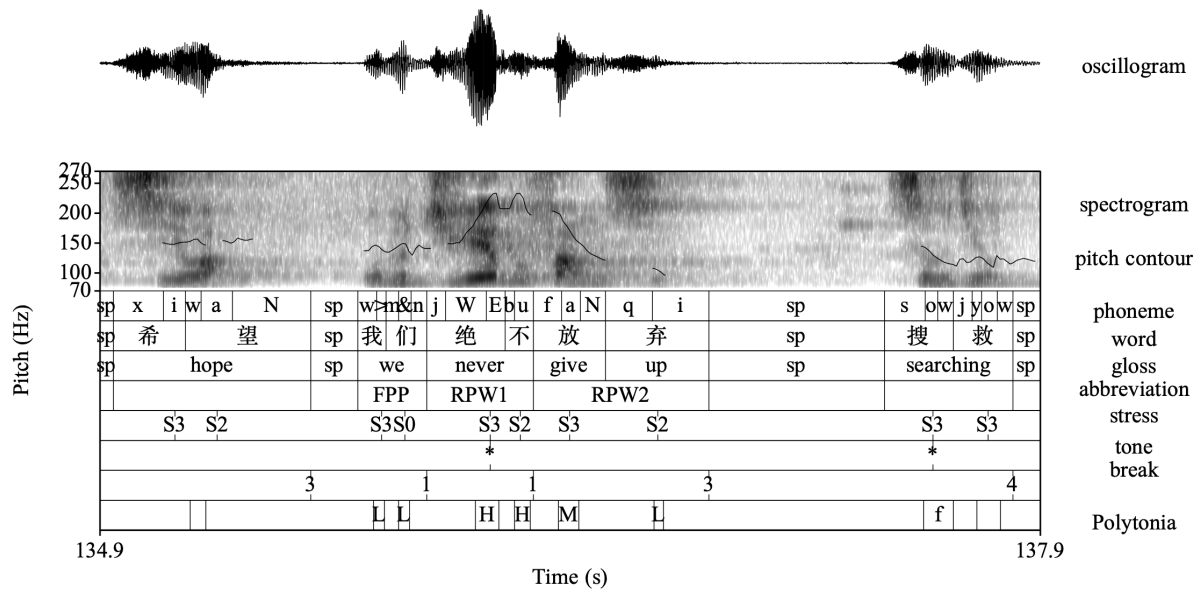
**Figure 1.** Semi-automatic analysis of an original Chinese segment

Figure 1 shows the FPP *wǒmen* to bear a low pitch (labelled L by Polytonia), whereas RPW1 *juébù* was delivered at the top of the speaker's pitch range (the top dashed line indicates the maximum) and was labelled H (high). The first syllable of RPW1 *jué* shows a clear pitch range expansion, with its host prosodic phrase probably terminated by the lengthened syllable *qì* of RPW2. Thus, semi-automatic analysis suggests the nucleus of this phrase to be in RPW1.

Auditory analyses were conducted to double-check the results of the semi-automatic transcription. These involved careful inspection of the oscillogram, spectrogram, and  $F_0$  contour obtained with Praat (Boersma & Weenink 2020), supplemented by repeated listening. Figure 2 illustrates the results for the segment in Figure 1, confirming the nucleus of the phrase to be in RPW1. Auditory analyses revealed that neither of the second syllables of the FPP (*men*)



or RPW1 (*bù*) showed extreme phonetic loss to justify labelling as 0, and FPP and RPW1 BREAK were accordingly tagged 1.



**Figure 2.** Auditory and instrumental analyses with Praat

I conducted prosodic annotation with a two-month interval. The intra-annotator agreement rates between the two transcriptions average 93.7% for OC, 89.23% for IE, and 94.42% for NE. The relatively low agreement rates for IE may result from the lack of pitch variation in the interpreted speech (Lenglet & Michaux 2020; Liu 2020), which made identifying pitch peaks difficult.

### 4.3 Statistical analyses

Statistical analyses were performed with R 4.0.3 (R Core Team 2020). A series of binary logistic regressions were first conducted to test and validate the plural preference claim. The performance of the models was measured with goodness-of-fit statistics, classification accuracy, prediction accuracy by means of the *C*-score, and Nagelkerke's  $R^2$ . The classification accuracy was compared to the no-information-rate baseline, which always chooses the more frequent level of the response. The *C*-score ranges from 0.5 to 1, where 0.5 represents random guessing and 1 refers to the theoretical maximum of prediction accuracy (Levshina 2015).

The MuPDAR(F) approach<sup>5</sup> was then used to test Hypothesis 1 and part of Hypothesis 2 about complex forms in production. MuPDAR(F) comprises two steps: (i) identifying the social, cognitive, and linguistic conventions in the FPP subject realisation of the source and recipient languages (i.e. native Mandarin and English) in political institutional settings, and (ii) measuring IE's deviation from such conventions.

First, two generalised linear mixed-effects models (glmm) were fitted on OC and NE separately to predict the FPP FORM; these are referred to as  $R_{OC}$  and  $R_{NE}$  and they are considered to be the points of reference for IE. (Note that referent-related variables were excluded from these analyses: by definition, the singular 'I' always points to a single level of BONDEDNESS [non-NP] and GROUP [exclusive]; this is an example of complete separation.)

Most of the predictors in Table 2 have more than two levels and so multicollinearity was checked with generalised variance-inflation factors corrected for the degrees of freedom (corrected GVIFs; Levshina 2015). Models and variables were selected using likelihood ratio tests with a significance threshold of  $p = 0.05$  for random-effects terms and  $p = 0.1$  for fixed-effects terms, given the exploratory nature of this study (Gries 2021). Two  $R^2$ s are reported for glmm, namely,  $R^2_{marginal}$  for fixed effects only and  $R^2_{conditional}$  for both fixed and random effects.

Second,  $R_{OC}$  and  $R_{NE}$  were applied in the prediction and deviation analyses. The underlying idea is that if Hypothesis 1b holds, then the interpreters' plural in ideologically salient contexts should deviate from that in OC. And if Hypothesis 2b holds, the effects of cognitive processing on FPP choices in IE should deviate from those on OC and NE. The deviation represents the difference between the predicted choice by the source or recipient language models ( $R_{OC}$  and  $R_{NE}$ ) and the interpreter's actual choice. The predicted choice is obtained by separately applying the coefficients of the fixed-effects variables of  $R_{OC}$  and  $R_{NE}$  to predict the FPP choice in IE. Because IE and NE feature different content, TOPIC was removed from  $R_{NE}$  – the new model being denoted  $R_{NE'}$  – before it was used. The deviation is obtained by comparing the predicted and actual choices and it is operationalised in two ways:

---

<sup>5</sup> A reviewer pointed out that the analyses performed using the MuPDAR(F) can alternatively be conducted by allowing VARIETY to interact with all non-source, non-referent, fixed-effect variables to predict the FPP choice. When the interaction method was applied to the OC and IE dataset, multicollinearity issues led to the removal of TOPIC, and therefore the confirmatory goal of the present study could not be achieved. The best-fit interaction models and the MuPDAR(F) reveal convergent findings that support Hypothesis 2. Analyses of the interaction models are given in Appendix A.

the logical variable SOURCE/RECIPIENT-LIKE and the numeric DEVIATION SCORE. SOURCE/RECIPIENT-LIKE measures whether the interpreter's actual FPP choice is identical to the predicted choice; DEVIATION SCORE quantifies the direction of the differences. The deviation is structured in such a way in MuPDAR(F) that:

- SOURCE/RECIPIENT-LIKE is true and DEVIATION SCORE is 0 when the interpreter made a source or recipient language-like choice.
- SOURCE/RECIPIENT-LIKE is false and DEVIATION SCORE  $> 0$  when the interpreter used the plural where a Mandarin or an NE speaker would use the singular.
- SOURCE/RECIPIENT-LIKE is false and DEVIATION SCORE  $< 0$  when the interpreter used the singular where a Mandarin or an NE speaker would use the plural.

The models of the deviation analyses are referred to as  $R_{OC-IE}$  and  $R_{NE'-IE}$ . Each model comprises a glmm with SOURCE/RECIPIENT-LIKE as the response and a linear mixed-effects model (lmm) with DEVIATION SCORE as the response. The glmms had the same predictors as the respective  $R_{OC}$  and  $R_{NE'}$  and underwent the selection processes described above. The lmm used the glmms' predictors.

The MuPDAR(F) approach outlined above does not deal with the ways in which institutional alignment (Hypothesis 1a), complex forms in the source and complex referents (Hypothesis 2), and zero-subject inputs (Hypothesis 3) have a bearing on the interpreters' preference for the plural. Therefore, two more regressions were fitted onto the plural cases ( $R_{plural}$ ) and the IE cases ( $R_{IE}$ ).  $R_{plural}$  targets Hypothesis 2 in the case of the participation of referent-related variables. By using MEDIATION STATUS as the response, it examines the effect of interpreting on plurals ('interpretational effects').  $R_{plural}$  considers only fixed effects because the random variables completely separate the response. Conversely, the glmm  $R_{IE}$  brings source-target relationships into the equation to predict the interpreters' FPP FORM ('contrastive effects') relating to Hypotheses 1a, 2a, and 3.  $R_{IE}$  includes:

- (i) predictors of  $R_{OC}$ , VOICE (section 2.1), and OTHER-PRIMING annotated on corresponding source segments of IE, which index the source context; and
- (ii) the non-referent variables shown in Table 2 measured in IE cases, which index the target context.

Crucially,  $R_{IE}$  tests Hypothesis 1a by allowing TOPIC to interact with all the fixed factors to determine which variables had topic-specific effects.

## 5. Findings

### 5.1 *Monofactorial models*

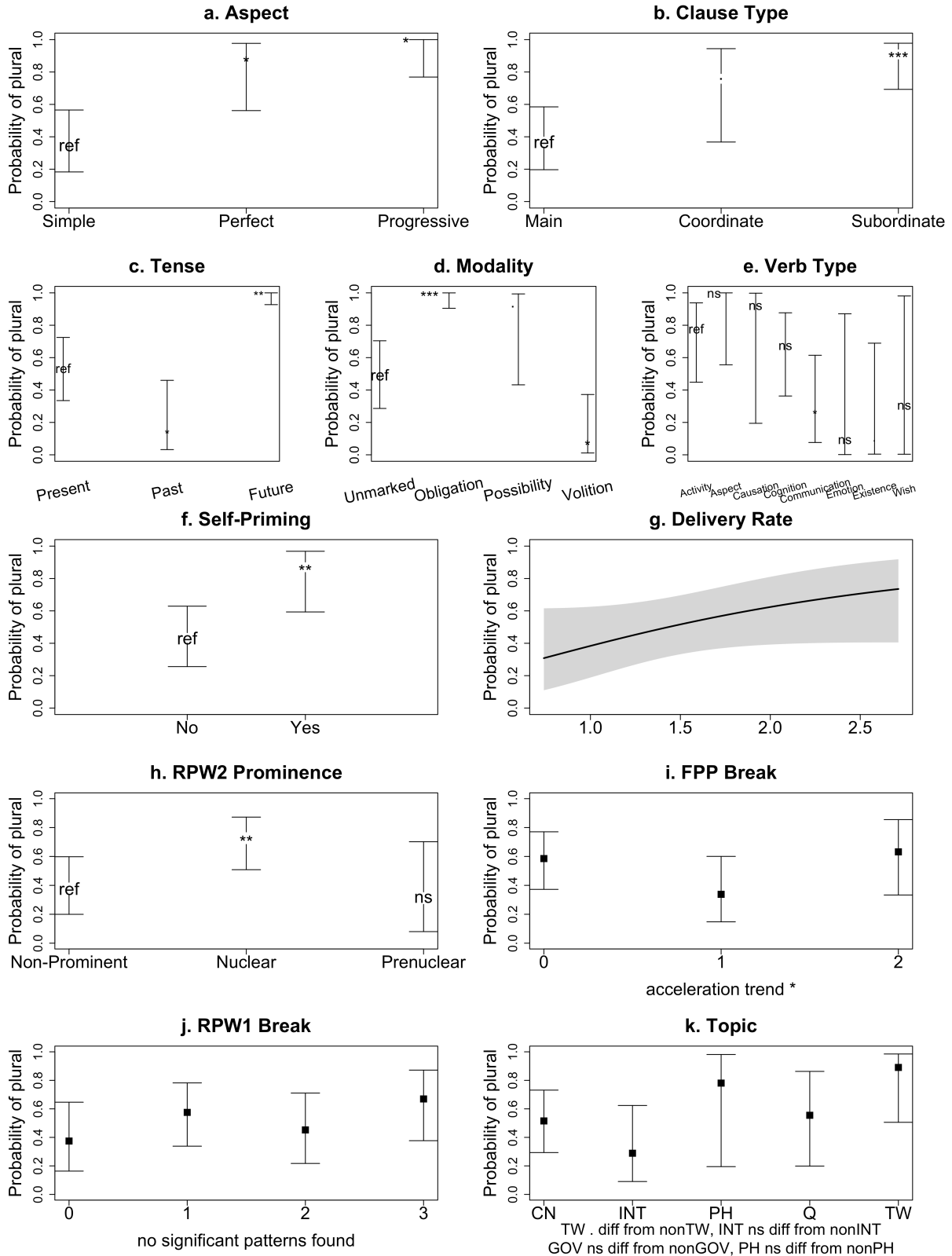
This section presents the probability that the plural will occur as a function of VARIETY. Given the difference in length between the IE and NE analysis samples (Table 1), the former was truncated to match the length of the latter in intra-lingual comparisons.  $M_{1a}$  is fitted on the OC and the IE and  $M_{1b}$  is fitted on the truncated IE and NE of the analysis samples.  $M_2$  relates to validation samples and is a replication of  $M_{1a}$  and  $M_{1b}$  using different samples. The full statistical reports are given in Appendix B.

$M_{1a}$  returns a highly significant but very weak fit between VARIETY and FPP FORM ( $p < 0.001$ , Nagelkerke's  $R^2 = 0.025$ ). It predicts a higher probability of the plural in IE relative to OC ( $\beta = 0.567$ ), lending statistical support to the claim that the plural is preferred by the interpreters. The classification accuracy of  $M_{1a}$  is 57.29%, which is significantly higher than the no-information-rate baseline of 52.73% ( $p_{binomial} < 0.001$ ). With a  $C$ -statistic of 0.566, the prediction accuracy of this monofactorial model is just slightly better than that of random guessing. Concurrently,  $M_{1b}$  demonstrates a lower probability of the plural being used in IE compared to NE ( $\beta = -0.896$ ). This means that the English native speaker showed a greater plural realisation rate compared to the interpreters working into English B. Coupled with the findings about the plural preference in both English–German and German–English SI in the EP (Beaton-Thome 2010), the result of  $M_{1b}$  excludes the possibility of linguistic proficiency determining the plural preference. According to  $M_2$ , in the validation samples the plural is also preferred by the interpreters in relation to the source speakers and by President Obama vis-à-vis the interpreters. The plural preference of the interpreters therefore holds for both the analysis and the validation samples. It is accordingly relatively safe to proceed with the following analyses.

### 5.2 *Multifactorial models of source/recipient languages*

#### 5.2.1 Exploring FPP choices of OC speakers: $R_{OC}$

The analyses of  $R_{OC}$  included 632 cases. The final, minimally adequate  $R_{OC}$  includes 11 variables. No multicollinearity (corrected GVIFs < 1.45), residual uniformity, dispersion, or outlier issues were found (see the diagnostic plots in Figure C1).



Note. \*\*\*  $p < 0.001$ , \*\*  $p \leq 0.01$ , \*  $p < 0.05$ , .  $p < 0.1$

**Figure 3.** Singular versus plural in OC;  $p$  values of less than 0.1 are summarised with a period.

The model's classification accuracy is very high (97.63%; cf. the no-information-rate baseline at 43.99%,  $p_{\text{binomial}} < 10^{-16}$ ). With a  $C$ -index of 0.998, the predictive power of this model is very close to the theoretical maximum of 1.  $R^2_{\text{marginal}}$  and  $R^2_{\text{conditional}}$  are respectively 0.607 and 0.897, encouragingly showing that fixed effects 'did more work' than random effects. Figure 3 shows the effect plots and the statistical report is given in Table C1. In each effect plot, a greater predicted probability value ( $y$ -axis) indicates an increased likelihood of the plural being used for the given value of that predictor ( $x$ -axis) when all the other predictors in the model are held constant. Apart from TOPIC, each level of the categorical variables is compared to the reference level (ref).

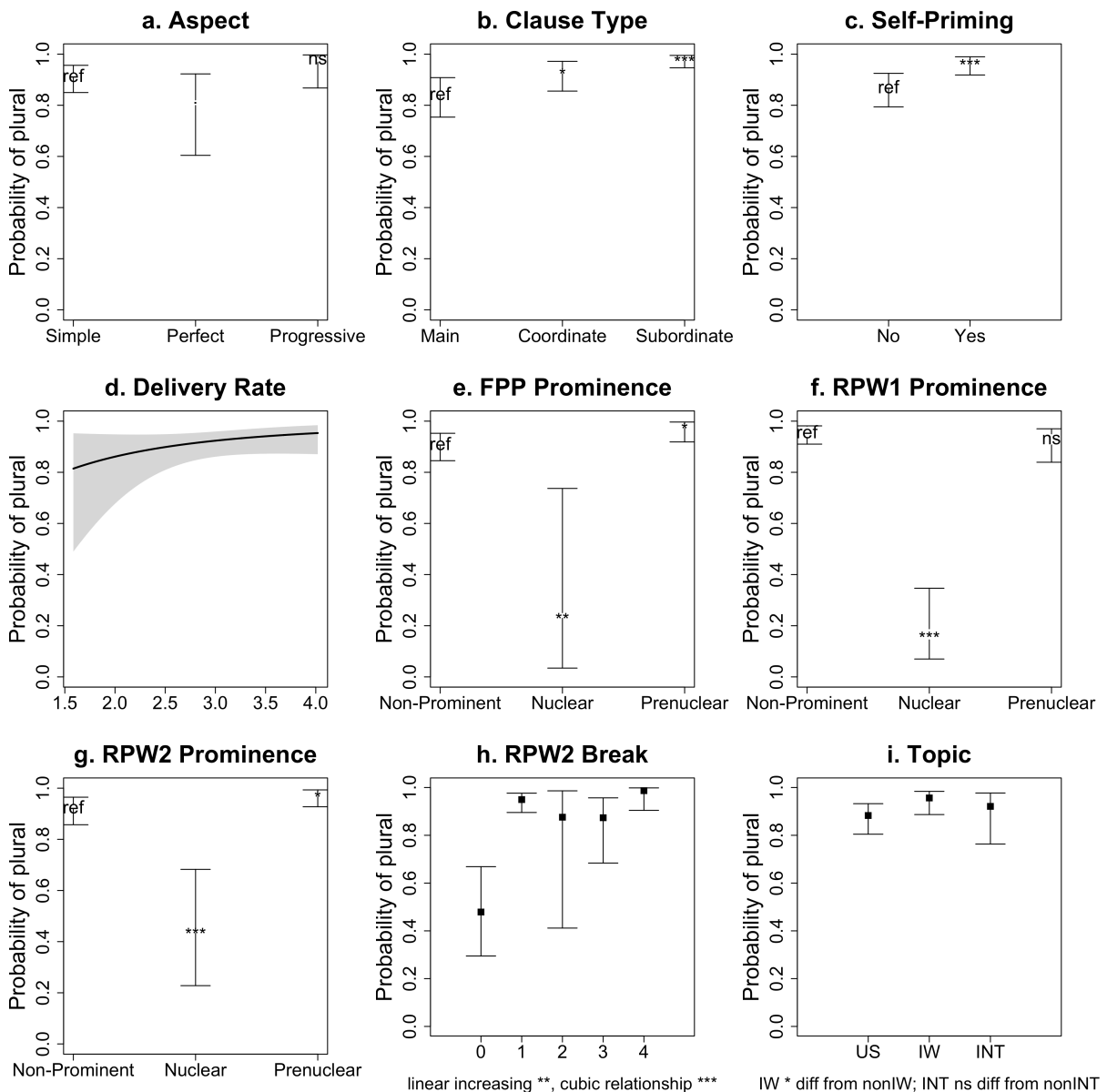
The results support Hypothesis 1a to a marginal extent and Hypothesis 2a. Figures 3a and 3b show significant differences in the predicted choice between the singular and the plural when aspect is simple and when it is perfect or progressive, and also when clause type is main and when it is coordinate or subordinate. This means that complex morphology (perfect and progressive) and syntax (coordinate and subordinate) favoured the plural. Figures 3c–3e show the higher predicted probability of the plural in contexts of future-tense auxiliaries, obligation and possibility modals, aspect verbs, and causation verbs relative to the present tense, unmarked modality, and activity verbs. Priming and faster delivery rates increase the odds of the plural being used (Figures 3f and 3g). Figure 3h shows the high probability of the plural's RPW2 being the nucleus rather than being non-prominent. Because a prosodic phrase has only one nucleus, neither the plural nor its RPW1 probably received primary stress. Figure 3i shows that if the FPP had undergone segmental loss, then it is probably the plural (FPP break = 0; note the large confidence interval associated with the effect of FPP break = 2). No significant effect of RPW1 break on FPP choices has been identified (Figure 3j).

Figure 3k demonstrates that the effects of social explanations on FPP choices in OC were weak. Relative to mainland Chinese and international matters, the topic of Taiwan raises the odds of the plural being used by Chinese premiers to a marginally significant extent ( $p = 0.061$ ); and the effect of the topic of international relations on the predicted probability of the plural does not reach statistical significance ( $p = 0.206$ ; as in Example 1). There was a trend of the plural being preferred in housekeeping remarks rather than in other topics, but the data

sparsity in them (see the large confidence interval) makes it difficult to draw reliable conclusions.

### 5.2.2 Exploring FPP choices of the NE speaker: $R_{NE}$

An exploration of the random effects revealed that the NE singular and plural had distinct MATCH and LEMMA preferences. A lexical bias plot of the FPPs is given as Figure C2. It shows that modal verbs and aspect, causation, and cognition verbs (e.g. ‘make’, ‘continue’, ‘trust’) co-occurred almost exclusively with the plural. The biases are broadly similar to those of the OC plural (Figures 3d and 3e). LEMMA, MATCH, MODALITY and VERB TYPE therefore cause complete separation and were removed from the analyses.  $R_{NE}$  considered 739 cases.



Note. \*\*\*  $p < 0.001$ , \*\*  $p \leq 0.01$ , \*  $p < 0.05$ , .  $p < 0.1$

**Figure 4.** Singular versus plural in NE;  $p$  values of less than 0.1 are summarised with a period.

The final  $R_{NE}$  includes nine variables. No red flags of multicollinearity (corrected GVIFs < 1.50), uniformity, dispersion, or outliers were raised (see Figure C3). The model's classification accuracy is 93.37% (cf. the no-information-rate baseline at 71.77%,  $p_{binomial} < 10^{-16}$ ), with a  $C$ -index of 0.970.  $R^2_{marginal}$  and  $R^2_{conditional}$  are 0.556 and 0.687 respectively. The statistical report is provided in Table C2 and effect plots in Figure 4.

The results support Hypotheses 1a and 2a. A comparison of Figures 4a and 3a shows the weaker effects of complex perfect and progressive aspect on raising the odds of the plural being used in NE than in OC; this was possibly owing to the high realisation rate of the plural in the NE dataset overall (0.75). Figure 4b shows a pattern consistent with that in Figure 3b: complex syntax favoured the plural in both OC and NE. Apart from the rise in the co-occurrence frequency (see above), the predicted probability of the plural is high when priming occurs (Figure 4c) and the president's delivery rate increases (Figure 4d).

Regarding the prosodic parameters, Figures 4e–4g demonstrate significant differences in the predicted probability of the plural being used when the FPP, RPW1, or RPW2 is unstressed and when it is the nucleus: overall, plural constructions in NE were unlikely to have attracted nuclear prominence – a sign of phonetic erosion. There was some likelihood of the plural receiving prenuclear, secondary accentuation compared to its being non-prominent (Figure 4e). Furthermore, Figure 4h shows that when the intonational phrase is terminated (RPW2 BREAK = 4) or continued in fluency following RPW2 (RPW2 BREAK = 1), the preceding FPP is likely to be the plural. In cases of phrase termination, the plural's RPW2 was probably the nucleus and the plural itself might have attracted prenuclear stress, as seen in Figures 4e and 4g. In cases of continued fluency, the plural construction was integrated with the next intonational phrase, signifying a loss of phonological independence.

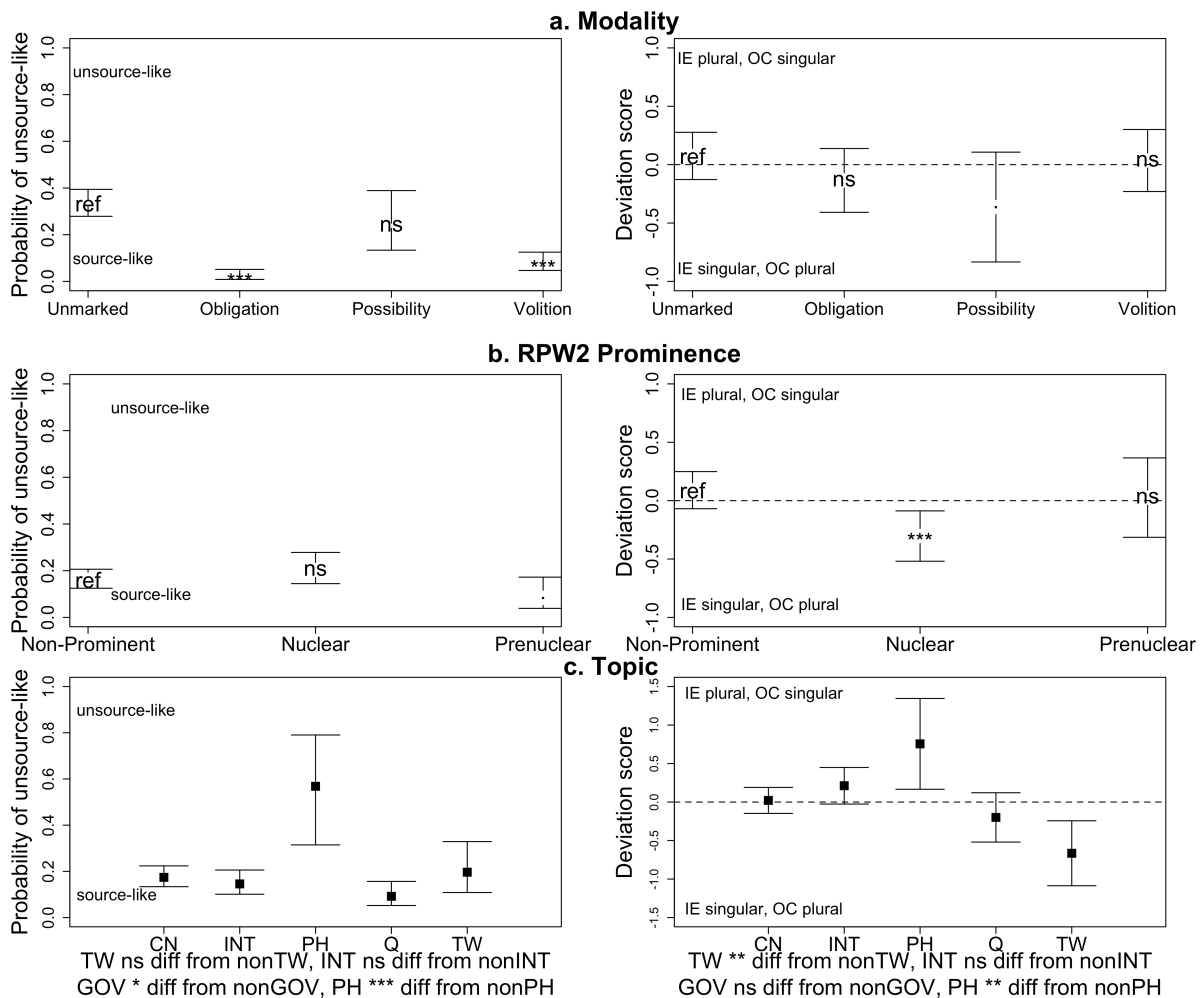
Regarding institutional alignment (Figure 4i), the predicted probability of the plural is high when the topic is the Iraq War ( $p = 0.021$ ), but international issues did not raise the likelihood of the plural being used in contrast to the trend with non-international topics ( $p = 0.474$ ).



### 5.3 Comparisons between interpreting and source/recipient-language conventions

#### 5.3.1 Applying $ROC$ to predict interpreter FPP choices: $ROC_{IE}$

$ROC_{IE}$  examines (i) whether the interpreters made the FPP choice that the source speakers would have made in a similar context and (ii) the direction of the deviation. The analyses comprised 1,042 IE cases.  $ROC$  predicts 76.68% of interpreter choices in such cases correctly, indicating a high degree of similarity in the FPP choices between OC and IE. The final model includes three variables (see Tables D1 and D2). Figure 5 shows the effect plots, with those of SOURCE-LIKE in the panels on the left and those of DEVIATION SCORE on the right.



Note. \*\*\*  $p < 0.001$ , \*\*  $p \leq 0.01$ , \*  $p < 0.05$ , .  $p < 0.1$

**Figure 5.** Deviation in FPP choices between OC and IE;  $p$  values of less than 0.1 are summarised with a period.

The results do not support Hypothesis 1b but confirm Hypothesis 2b. Figure 5a shows that if an FPP case uses obligation modals, then the interpreters will probably make the source-like

choice of the plural ( $\beta = -3.153$ ; see Figure 3d); and if the case involves volition modals, the interpreters will prefer the source-like singular ( $\beta = -1.787$ ). The difference in the regression coefficients ( $-3.153$  versus  $-1.787$ ) reveals a stronger effect of obligation modals on the interpreters' source-like choice of the plural than that of volition modals on the interpreters' singular. This signifies the more frequent co-occurrence of the former chunk than the latter.

The right panel of Figure 5b shows that when the RPW2 is the nucleus, the preceding FPP is more likely to be the singular in IE and the plural in OC. In other words, the IE plural's RPW2 is unlikely to be the nucleus – a sign of stress loss. The absence of the effect of FPP BREAK on the deviation is telling: similarly to that in OC (Figure 3i), if the FPP had undergone phonetic erosion, then it is probably the plural in IE.

The patterns identified in Figure 5c do not support Hypothesis 1b but confirm the role of situational settings in explaining the plural preference in a small number of IE cases.

First, the right panel shows that, compared to the matters of mainland China and international relations, the interpreters used the singular on the topic of Taiwan, whereas Chinese premiers favoured the plural (Figure 3k). This means that the interpreters' plural on this ideologically salient topic *decreased* in frequency vis-à-vis that of the source speakers. No significant deviation in FPP choices between OC and IE is found on international relations compared to the topics of mainland China and Taiwan. The deviation concerning interpreting for premiers versus reporters stems from the relative backgrounding of 'we' on Taiwan.

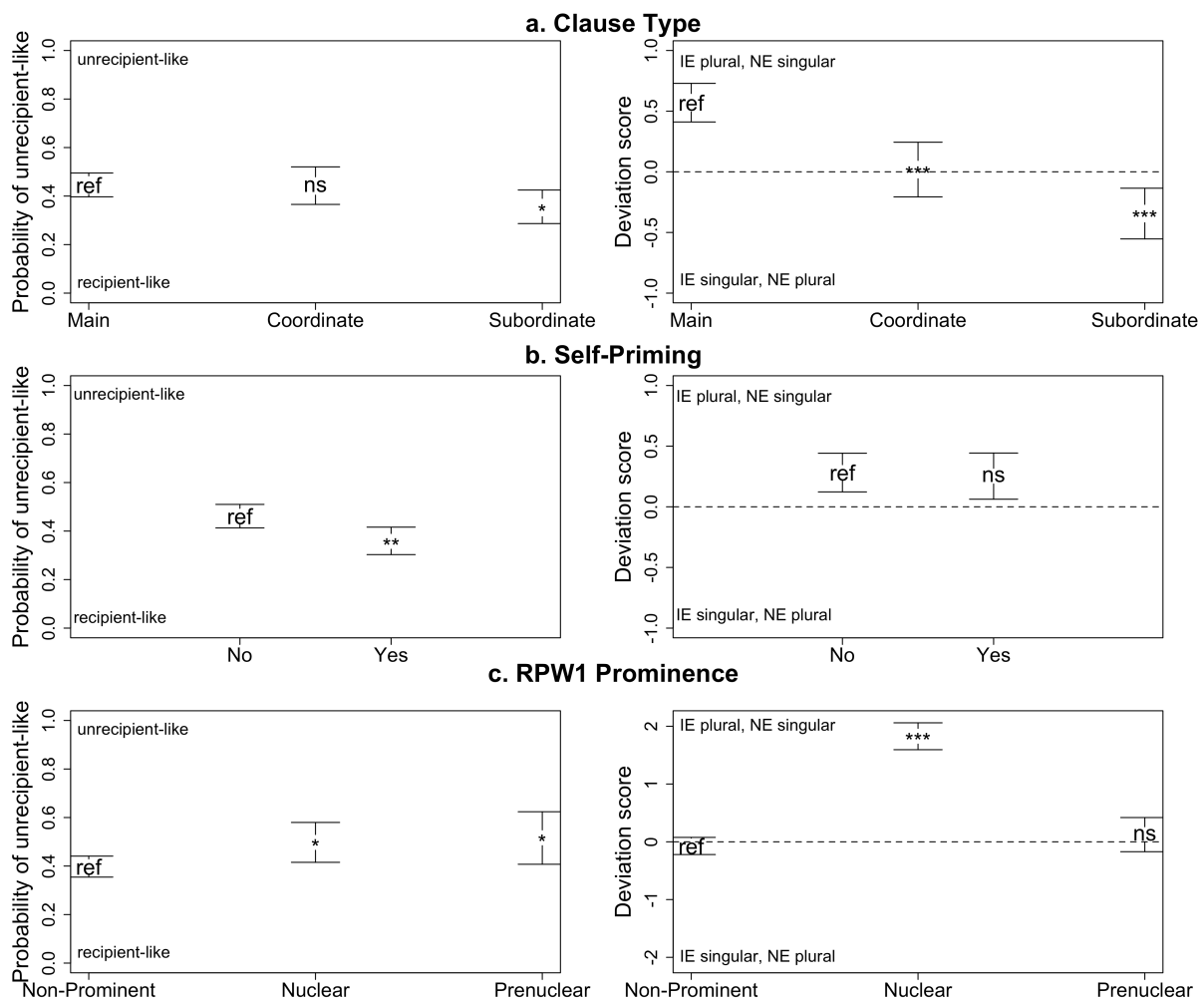
Second, in relation to non-procedural remarks, the plural was favoured on the subject of housekeeping to a significantly greater degree in IE than in OC, but the confidence bands are rather large, indicating a sparsity of data.

### 5.3.2 Applying $R_{NE'}$ to predict interpreter FPP choices: $R_{NE'-IE}$

$R_{NE'-IE}$  examines (i) whether the interpreters made the FPP choice that the NE-speaker would have made in a similar context and (ii) the direction of the deviation. The analyses considered 1,005 IE cases. The prediction accuracy of  $R_{NE'}$  on IE data is 56.22% and the final  $R_{NE'-IE}$  includes three fixed-effect variables (see Tables D3 and D4).

The results lend qualified support to Hypothesis 2. The left panel of Figure 6a shows that, compared to the main clause, subordinate clauses increase the odds of the interpreters' recipient-like choice of the plural (Figure 4b). The right panel further indicates that in cases involving main clauses, the interpreters would use the plural whereas the NE-speaker used the singular. An inspection of the underlying data reveals that, overall, the interpreters preferred

main clauses over coordinate and subordinate clauses to a greater extent than the NE-speaker. This indicates that the interpreters also avoided contexts of grammatical and cognitive complexity.



Note. \*\*\*  $p < 0.001$ , \*\*  $p \leq 0.01$ , \*  $p < 0.05$

**Figure 6.** Deviation in FPP choices between NE and IE

Concomitantly, priming raises the likelihood of the interpreters' recipient-like choice of the plural (Figures 6c and 4c). The left panel of Figure 6c shows that the probability of unrecipient-like choices increases when RPW1 is nuclear or prenuclear; the right panel indicates that the interpreters placed the plural's RPW1 under primary stress whereas the NE-speaker would produce the singular's RPW1 as the nucleus. In that case, the IE plural is unlikely to be the nucleus.

## 5.4 Exploring institutional interpreters' plural preference

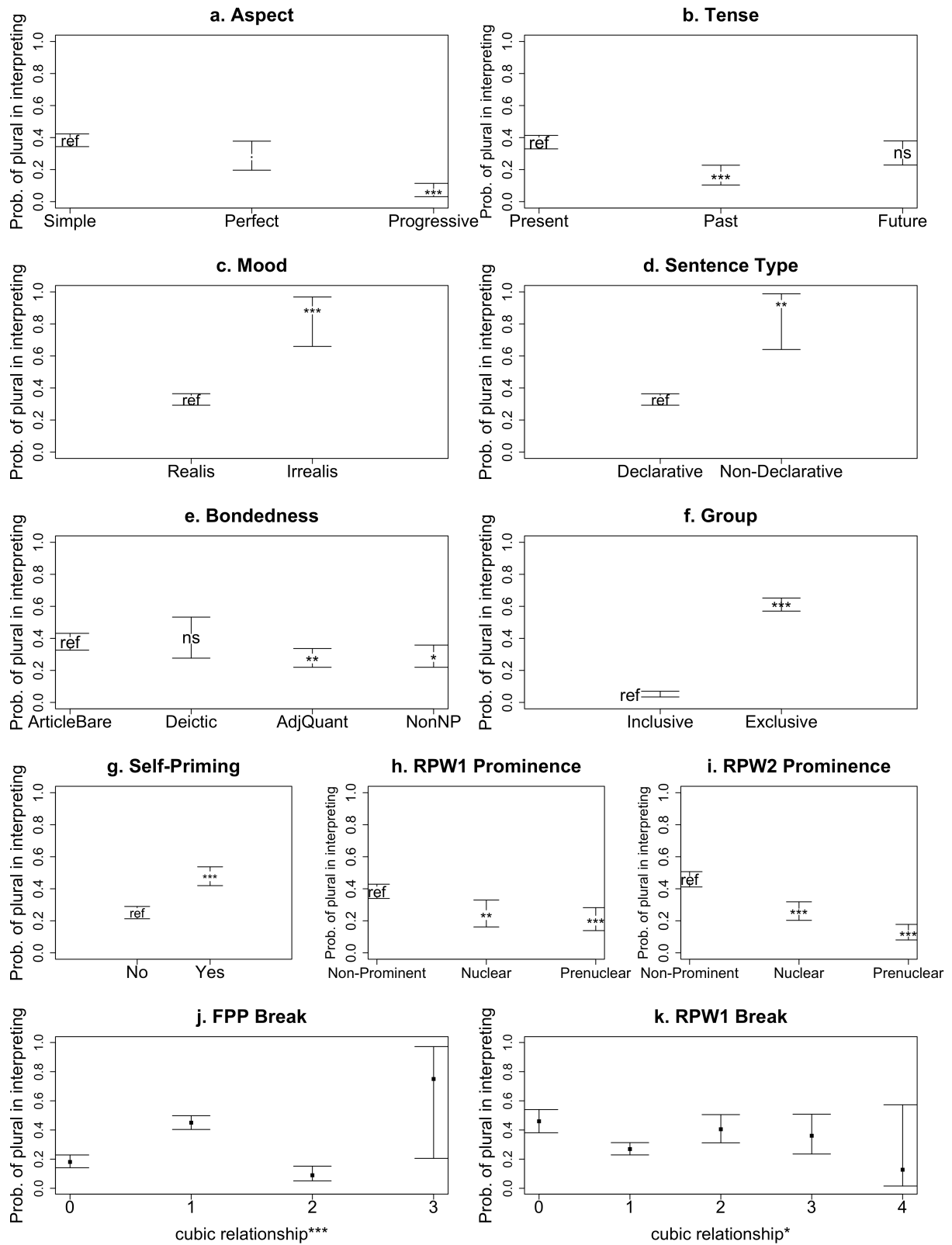
### 5.4.1 Interpretational effects: $R_{plural}$

The analyses considered 1,450 cases. The final model includes 11 variables, and no problems regarding multicollinearity (corrected GVIFs < 1.21), uniformity, dispersion, or outliers were found (Figure E1).  $R_{plural}$  predicts the MEDIATION STATUS of 82.69% cases correctly (cf. the no-information-rate baseline at 42.07%,  $p_{binomial} < 10^{-16}$ ), with a reassuring  $C$ -statistic of 0.892 and Nagelkerke's  $R^2$  at 0.560. The effect plots are in Figure 7 and the statistical report is in Table E1.

The results confirm Hypothesis 2. Regarding the plural forms, Figures 7a and 7b show the high predicted probability of the IE plural when ASPECT is simple and when TENSE is present or future. The patterning demonstrates that, compared to the non-interpreters (i.e. premiers, reporters, moderators, and the US President), the interpreters favoured the plural in simple morphology and with future-tense auxiliaries for the sake of processing ease. Figures 7c and 7d indicate that if the plural occurs in the complex morphosyntax (irrealis mood and a non-declarative sentence), it has probably been produced by the interpreters.

Regarding the plural's underlying referents, Figure 7e shows significant differences in the predicted probability of the IE plural when the referent is bare or modified by articles and when it is modified by adjectives, quantifiers, or non-NP. This indicates a positive relationship between the bondedness of referents and the odds of the plural in interpreting. The patterns mean that the interpreters had fine-tuned their referent preferences to the bondedness hierarchy. Institutional alignment cannot explain such hierarchical effects. Figure 7f shows that, relative to inclusive referents, more bonded exclusive referents increase the probability of the plural in IE rather than in OC or NE.

Priming (Figure 7g) and phonetic erosion (Figures 7h–7k) increase the odds of the plural in IE. The absence of the effect of FPP PROMINENCE in differentiating the plural in interpreting from that in non-interpreting is striking. This suggests that, compatible with the plural's de-accentuation in NE (Figures 4e and 6c) and OC (Figures 3h and 5b), the interpreters' plural was unlikely the nucleus, with its RPW1 and RPW2 probably not being prominent (Figures 7h and 7i). If the plural is fluently delivered (FPP BREAK = 1; Figure 7j) and its RPW1 and RPW2 are reduced (RPW1 BREAK = 0; Figure 7k), then it is probably produced by the interpreters.



Note. \*\*\*  $p < 0.001$ , \*\*  $p \leq 0.01$ , \*  $p < 0.05$ , .  $p < 0.1$

**Figure 7.** The plural in interpreting versus non-interpreting;  $p$  values of less than 0.1 are summarised with a period.

### 5.4.2 Contrastive effects: $R_{IE}$

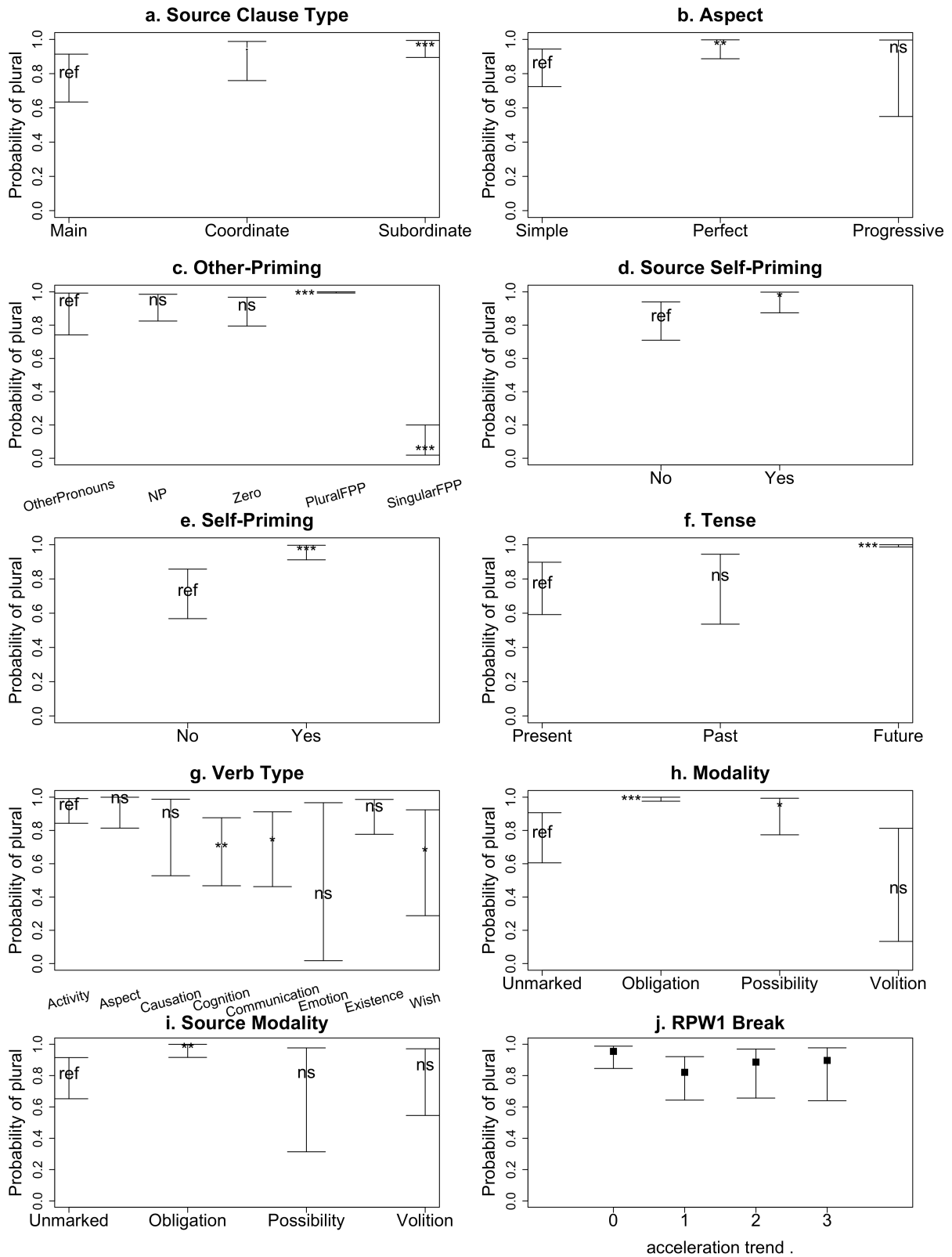
For  $R_{IE}$ , data sparsity forced source VERB TYPE levels to be conflated into ActAspCau (activity, aspect, and causation), cognition, communication, emotion-wish, and existence. The housekeeping cases had to be discarded. As in NE, LEMMA and MATCH completely separate the FPP choice (see Figure E2) and were therefore removed. The analyses considered 1,022 cases.

The final model returns ten fixed predictors and an interaction term between TOPIC and source VERB TYPE (Table E2). Multicollinearity was not an issue (corrected GVIFs < 1.79). No red flags were raised concerning residuals (Figure E3).  $R_{IE}$  arrived at a classification accuracy of 96.97% (cf. the no-information-rate baseline at 58.12%,  $p_{binomial} < 10^{-16}$ ), with the  $C$ -value,  $R_{marginal}^2$ , and  $R_{conditional}^2$  at 0.995, 0.836 and 0.931 respectively. The effect plots of fixed factors are given in Figure 8; those of the interaction term are shown in Figure 9.

The results corroborate Hypotheses 2a and 3 and largely discredit Hypothesis 1a. Figures 8a and 8b show the high probability of the plural in contexts where the source syntax and the target morphology are complex relative to simple clauses and aspect. Figure 8c indicates that the odds of interpreter uptake of ‘we’ are high in response to non-FPPs, complex NPs, zero subjects in Mandarin, and plural FPPs.

Hypotheses 2a and 3 are further supported by Figures 8d–8j. Self-priming in the source ( $\beta = 2.383$ ; Figures 8d and 3f) and target ( $\beta = 2.986$ ; Figure 8e) raise the odds of the plural rather than the singular being used, the combined effects of which ( $2.383 + 2.986 = 5.369$ ) outweigh those of the plural lemma ( $\beta = 4.277$ ; Figure 8c). The other-priming of the plural lemma was weaker than that of the singular on the interpreters’ FPP choices ( $\beta = -5.656$ ; cf. Li and Wang [2012]).

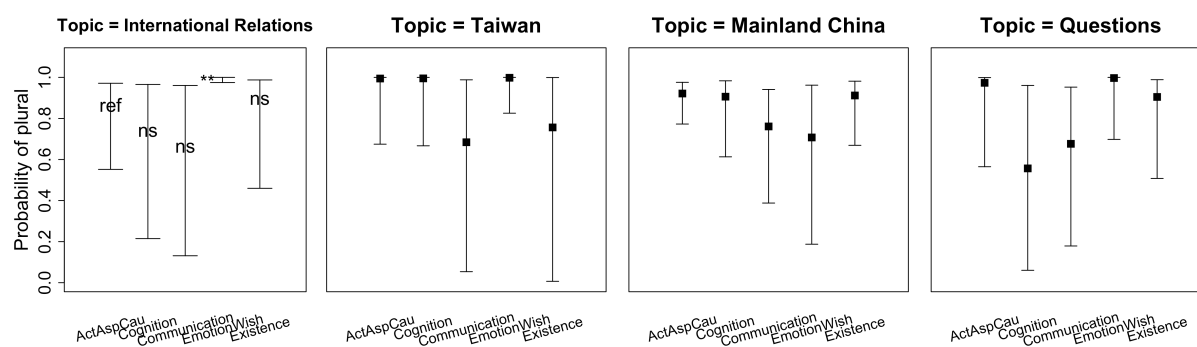
In contrast to the singular, the plural shows a pervasive increase in the frequency of co-occurrence with future-tense auxiliaries (Figure 8f) and verbs denoting activity, aspect, causation, and existence (Figure 8g). The plural also tends to co-occur with obligation and possibility modals (Figure 8h), which is probably other-primed by the obligation modals in the source (Figures 8i and 3d). Phonological boundary erasure likely occurred in IE plural constructions (RPW1 BREAK = 0; Figure 8j). Such patterning indicates that the whole plural construction rather than the single lemma represented a meaningful production unit for the interpreters – a sign of holistic storage and retrieval in cognition.



Note. \*\*\*  $p < 0.001$ , \*\*  $p \leq 0.01$ , \*  $p < 0.05$ , .  $p < 0.1$

**Figure 8.** Singular versus plural in IE;  $p$  values of less than 0.1 are summarised with a period.

When language production is treated beyond the single-word level, instances of FPP shifts such as the optional shift in Example 1 receive a unified explanation that considers the social, cognitive, and linguistic factors together. Figure 8g shows that, compared to activity verbs, emotion- and wish-class verbs such as ‘hope’ were unlikely pairs with the plural in IE, but other-priming from such predicates in the source (Figure 9) overcame the disassociation – only when the topic concerned international relations.



Note. \*\*  $p \leq 0.01$

**Figure 9.** The effects of TOPIC : source VERB TYPE on FPP FORM in IE

With the caveat that not all of the relevant variables have been discovered, an instantiation of Toury’s (2004) probabilistic account of translational behaviour emerges: if a predicate denoting emotion or a wish is used by a government speaker on international relations, then institutional interpreters will probably choose ‘we’ over ‘I’. Questions about the specific mechanisms underlying such an effect – politeness, identification with historical or institutional identities, etc. – are beyond the scope of this study, and I can only hazard that the interpreters concerned have conventionalised the shift from *wǒ xīwàng* to ‘we hope’ in their cognitive, work, and educational environments.

As cognitive and linguistic explanations hold sway, it seems highly likely that they have overridden institutional alignment in accounting for the institutional interpreters’ FPP plural preference. The only shifts that were ‘exceptional’ to the institutional interpreters might relate to source emotion and wish predicates on the subject of international relations. Beyond that, the institutional interpreters behaved similarly to anyone endowed with the cognitive ability of chunking: in the face of complex language-specific environments, they relied on grammaticalised constructions for the sake of processing efficiency.



## 6. Discussion and conclusions

On the status of pronouns in political discourse, van Dijk (2004: 734) cautioned against the *carte blanche* approach of treating first-person plural pronouns as “ingroup and outgroup markers”, because “most ideological variation will be found at the levels of meanings”. In the present study, I have found little support for the proposal that political institutional interpreters’ use of ‘we’ is ideologically salient. Multifactorial analyses based on the maximalist usage-based theory reveal that plural constructions in the interpreted speech have developed into cognitive units and grammatical functions with little semantic value. Institutional interpreters make the plural form a preference in compliance with the conventions of political speech (section 5.1) and also when processing and producing complex forms and referents in the target (sections 5.3.2 and 5.4.1) and source, and zero-subject inputs (section 5.4.2).

The institutional alignment explanation proposed in previous studies fails to account for the extensive grammaticalisation of plural constructions in interpreted speech; this is evidenced by the substantial increase in the frequency of co-occurrence, association with priming and fluency, morphosyntactic bondedness, and phonetic erosion. This was shown for plural constructions in interpreting in comparison to those in source and comparable speeches where the singular was the reference (section 5.3), to those in non-interpreting (section 5.4.1), and to the interpreters’ singulars (section 5.4.2). Once a range of intersecting social, cognitive, and linguistic factors are assembled, institutional alignment or monofactorial paradigms have little explanatory power.

Mechanisms underlying first-person shifts include the domain-general cognitive process of chunking that is by no means specific to institutional interpreters. Across unrelated languages (section 5.2) and variable mediation status (sections 5.3 and 5.4), such a mechanism drives grammatical and lexical items with very similar meanings (plural FPPs and aspect, causation, modal verbs, future-tense auxiliaries) nested in cross-culturally comparable contexts (Liu, in press) and cognitively complex environments to enter into grammaticalisation. This process gives rise to plural chunks that are accessed holistically in human cognition.

The usage-based approach and correspondingly refined multivariate methods are thus relevant to interpreting research in two ways. Rather than dichotomising source or recipient conventions, they show plural constructions to exist on a continuum of association strengths, from the less bonded to those more schematically general across languages. The question then

becomes the extent to which interpreters (ab)use such commonalities to their processing benefit. Rather than compartmentalising social, cognitive, or linguistic explanations, they offer a unified account of many facets of the interpreter's language that were previously treated as disparate (except by Li & Halverson 2020).

On the multivariate account, the joint effects of emotion- and wish-class predicates in the source on the FPP choices of institutional interpreters in the context of international relations may be the only alignment-related shifts identified in this study. However, comparisons between institutional and freelance interpreters are needed to validate such a finding. Multifactorial designs are ideal for exploring FPP use in SI and interpreting into one's A versus B language. Analyses of interpreter-specific and temporal idiosyncrasies require larger corpora than those used in the present study. An interplay between corpus and experimental work is a next step.

The present research is not intended to undermine the discourse-analytical approach; nor does it preclude the value of this method in corpus exploration or interpreter education. What this research illustrates is that an interpreter's choice is simultaneously constrained by and manifested at many levels that problematise the notion of a unidimensional, *a priori* framed 'institutional interpreter'. If the objective of Translation and Interpreting Studies is to "construct a system of interconnected, mutually conditioning statements" (Toury 2004: 25), then interpreters must be recognised as the complex, agentive persons that they are.

**References**

- Beaton-Thome, M. (2010). Negotiating identities in the European Parliament: The role of simultaneous interpreting. In M. Baker, M. Olohan & M. Calzada-Pérez (Eds.), *Text and context: Essays on translation and interpreting in honour of Ian Mason*. Manchester: St. Jerome, 117–138.
- Beaton-Thome, M. (2013). What's in a word? Your enemy combatant is my refugee: The role of simultaneous interpreters in negotiating the lexis of Guantánamo in the European Parliament. *Journal of Language and Politics* 12 (3), 378–399.
- Beckman, M. E. & Elam, G. A. (1997). Guidelines for ToBI labelling (version 3.0). [https://www.ling.ohio-state.edu/research/phonetics/E\\_ToBI/](https://www.ling.ohio-state.edu/research/phonetics/E_ToBI/) (accessed 30 August 2022).
- Biber, D., Johansson, S., Leech, G. N., Conrad, S. & Finegan, E. (2021). *Grammar of spoken and written English*. Amsterdam: John Benjamins.
- Boersma, P. & Weenink, D. (2020). *Praat: Doing phonetics by computer*. Version: 6.1.31.
- Bybee, J. L. (2010). *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Chao, Y. R. (1968). *A grammar of spoken Chinese*. Berkeley, CA: University of California Press.
- Cole, J. (2015). Prosody in context: A review. *Language, Cognition and Neuroscience* 30 (1–2), 1–31.
- Croft, W. (2003). *Typology and universals* (2nd ed.). Cambridge: Cambridge University Press.
- De Sutter, G. & Lefer, M.-A. (2020). On the need for a new research agenda for corpus-based translation studies: A multi-methodological, multifactorial and interdisciplinary approach. *Perspectives* 28 (1), 1–23.
- Dehé, N. & Wichmann, A. (2010). Sentence-initial I think (that) and I believe (that): Prosodic evidence for uses as main clause, comment clause and discourse marker. *Studies in Language* 34 (1), 36–74.
- Fu, R. & Chen, J. (2019). Negotiating interpersonal relations in Chinese–English diplomatic

- interpreting. *Interpreting* 21 (1), 12–35.
- Gries, S. Th. (2010). Behavioral profiles: A fine-grained and quantitative approach in corpus-based lexical semantics. *The Mental Lexicon* 5 (3), 323–346.
- Gries, S. Th. (2021). (Generalized linear) mixed-effects modeling: A learner corpus example. *Language Learning* 71 (3), 757–798.
- Gries, S. Th. & Deshors, S. (2014). Using regressions to explore deviations between interlanguage and native language: Two suggestions. *Corpora* 9 (1), 109–136.
- Gu, C. & Tipton, R. (2020). (Re-)voicing Beijing’s discourse through self-referentiality: A corpus-based CDA analysis of government interpreters’ discursive mediation at China’s political press conferences (1998–2017). *Perspectives* 28 (3), 406–423.
- Harris, B. (1990). Norms in interpretation. *Target* 2 (1), 115–119.
- Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. Oxford: Oxford University Press.
- Kruger, H. & De Sutter, G. (2018). Alternations in contact and non-contact varieties: Reconceptualising that-omission in translated and non-translated English using the MuPDAR approach. *Translation, Cognition & Behavior* 1 (2), 251–290.
- Langacker, R. W. (1988). A usage-based model. In B. Rudzka-Ostyn (Ed.), *Topics in cognitive linguistics*. Amsterdam: John Benjamins, 127–161.
- Lenglet, C. & Michaux, C. (2020). The impact of simultaneous-interpreting prosody on comprehension: An experiment. *Interpreting* 22 (1), 1–34.
- Levshina, N. (2015). *How to do linguistics with R*. Amsterdam: John Benjamins.
- Li, D. & Wang, K. (2012). Hanying tongchuan zhong cihui moshi de yuliao ku kaocha [A corpus-based study on lexical patterns in simultaneous interpreting from Chinese into English]. *Xiandai Waiyu* 35 (4), 409–415.

- Li, Y. & Halverson, S. L. (2020). A corpus-based exploration into lexical bundles in interpreting. *Across Languages and Cultures* 21 (1), 1–22.
- Lin, P. M. S. (2018). *The prosody of formulaic sequences: A corpus and discourse approach*. London: Bloomsbury.
- Liu, N. (in press). Register shifts in political conference interpreting: A multidimensional analysis. In J. Pan, S. L. Halverson & J. Munday (Eds.), *Translating and interpreting political discourse: New trends and perspectives*. Leiden: Brill, 1–30.
- Liu, N. (2020). The effect of political conference interpreting: Triangulating cognitive pragmatics, speech corpora and audience perception. PhD dissertation, University of Hong Kong.
- Mertens, P. (2014). Polytonia: A system for the automatic transcription of tonal aspects in speech corpora. *Journal of Speech Sciences* 4 (2), 17–57.
- Monacelli, C. (2009). *Self-preservation in simultaneous interpreting*. Amsterdam: John Benjamins.
- Narrog, H. & Heine, B. (2021). *Grammaticalization*. Oxford: Oxford University Press.
- Peng, S., Chan, M. K. M., Tseng, C., Huang, T., Lee, O. J. & Beckman, M. E. (2005). Towards a pan-Mandarin system for prosodic transcription. In S.-A. Jun (Ed.), *Prosodic typology: The phonology of intonation and phrasing*. Oxford: Oxford University Press, 230–270.
- Pierrehumbert, J. & Hirschberg, J. B. (1990). The meaning of intonational contours in the interpretation of discourse. In P. R. Cohen, J. L. Morgan & M. E. Pollack (Eds.), *Intentions in communication*. Cambridge, MA: MIT Press, 271–311.
- Plevoets, K. & Defrancq, B. (2018). The cognitive load of interpreters in the European Parliament: A corpus-based study of predictors for the disfluency uh(m). *Interpreting* 20 (1), 1–28.

- Pöchhacker, F. (1995). Simultaneous interpreting: A functionalist perspective. *HERMES* (14), 31–53.
- R Core Team (2020). *R: A language and environment for statistical computing*. Vienna, Austria.
- Toury, G. (2004). Probabilistic explanations in translation studies: Welcome as they are, would they qualify as universals? In A. Mauranen & P. Kujamäki (Eds.), *Translation universals: Do they exist?* Amsterdam: John Benjamins, 15–32.
- van Dijk, T. A. (2004). Politics, ideology, and discourse. In R. Wodak (Ed.), *Encyclopedia of language and linguistics* (2nd ed.). Amsterdam: Elsevier, 728–739.
- Yang, C. (2016). *The acquisition of L2 Mandarin prosody*. Amsterdam: John Benjamins.
- Yuan, J. & Liberman, M. (2008). Speaker identification on the SCOTUS corpus. *The Journal of the Acoustical Society of America* 123 (5), 3878.