

UC Santa Barbara

UC Santa Barbara Previously Published Works

Title

The supralaryngeal articulation of stress and accent in Greek

Permalink

<https://escholarship.org/uc/item/21w2k7bk>

Authors

Katsika, Argyro

Tsai, Karen

Publication Date

2021-09-01

DOI

10.1016/j.wocn.2021.101085

Peer reviewed



HHS Public Access

Author manuscript

J Phon. Author manuscript; available in PMC 2022 September 21.

Published in final edited form as:

J Phon. 2021 September ; 88: . doi:10.1016/j.wocn.2021.101085.

The supralaryngeal articulation of stress and accent in Greek

Argyro Katsika^{a,b,*}, Karen Tsai^a

^aUniversity of California Santa Barbara, Department of Linguistics, South Hall 3513, Santa Barbara, CA 93106, USA

^bHaskins Laboratories, 300 George Street, Suite # 900, CT 06511, USA

Abstract

It is well reported that articulatory movements comprising prominence units are longer, larger and faster than their non-prominent counterparts. However, it is unclear whether these effects arise at the level of lexical stress or accent or both, reflecting a hierarchy of prominence, i.e., being stronger when induced by accent as opposed to stress. It is also uncertain whether prominence-induced kinematic effects are invariant across positions of stress within the word, types of focus the accent denotes, and positions of words in the phrase. We conduct an electromagnetic articulography (EMA) study to assess the supralaryngeal kinematic correlates of prominence in Greek across three stress positions (antepenultimate, penultimate, ultimate; i.e., all possible stress positions in Greek), two accentual conditions (accented and de-accented) and two phrasal positions (phrase-medial and phrase-final). Focus type is also considered, with the accentual conditions coming from two types of focus (broad and narrow), while the de-accented conditions are by default unfocused. Our results indicate that stressed syllables involve longer, larger and faster gestures than their unstressed counterparts, regardless of the position of stress within the word. Notably, variation in velocity is accounted for by variation in displacement. Presence of accent does not further expand the stressed gestures, although it is related to minimal kinematic changes across the whole word, the exact profile of which depends on stress position. With the exception of final vowel duration, focus type is not systematically encoded in these kinematic effects. Finally, interactions are detected between the kinematic profile of prominence and that of boundaries. Implications of our findings for the hierarchy of prominence and cross-linguistic differences are discussed, and a gestural account of prominence and boundaries is put forward.

Keywords

Stress; Pitch accent; Prominence; Strengthening; Lengthening; Articulation; Greek

*Corresponding author at: University of California Santa Barbara, Department of Linguistics, South Hall 3513, Santa Barbara, CA 93106, USA. akatsika@linguistics.ucsb.edu (A. Katsika).

Uncited references

Niebuhr (2010), Ogden (2012).

CRedit authorship contribution statement

Argyro Katsika: Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft, Writing - review & editing, Visualization, Supervision, Project administration, Funding acquisition. **Karen Tsai:** Formal analysis, Visualization, Validation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

1. Introduction

1.1. Towards a definition of prominence

A major goal of current theories of prosodic structure is to capture the grammatical nature of prosodic prominence. An important step towards this goal is to phonetically characterize the different degrees of prominence accurately. On the grammatical domain, prominence is considered one of the two functions – grouping is the other one – that language uses to organize its utterances in a hierarchical structure (Lieberman & Prince, 1977; Nespor & Vogel, 1986; Pierrehumbert & Beckman, 1988; Selkirk, 1984). On the phonetic domain, a large set of prominence correlates has been detected, with their presence or degree of contribution varying with language and prominence type (see Fletcher (2010) for an overview). The list includes, but is not limited to, the acoustic dimensions of duration, intensity, pitch height, formant patterns, and spectral tilt; and the articulatory dimensions of duration, displacement (spatial position), and velocity of articulatory movements. However, it is not clear whether, and if yes how, these phonetic dimensions reflect different degrees of prominence, and thus the number and order of hierarchical levels of prominence have yet to be specified. This would require a separation of the effects of lexical stress from those of phrasal prominence on the one hand, and among the different degrees of phrasal prominence on the other. To that end, articulatory data would be particularly valuable, since currently articulatory research on this topic is sparse, mainly derived from English, and primarily examining contrastive focus (e.g., Beckman, Edwards, & Fletcher, 1992; Beckman & Edwards, 1994; Cho, 2005, 2006; de Jong, 1991, 1995; de Jong, Beckman, & Edwards, 1993; Fowler, 1995; Harrington, Fletcher, & Beckman, 2000; Harrington, Fletcher, & Roberts, 1995). Recent articulatory research on German has taken important first steps towards considering other focus types in prominence marking (Mücke & Grice, 2014; Roessig & Mücke, 2019; see also Hermes, Becker, Mücke, Baumann, & Grice, 2008). Another piece of the puzzle that is missing is the effect of phrasal position on the manifestation of prominence, a point that becomes more relevant as recent findings point to an intricate relationship between the phonetic events of phrasal edges, such as boundary-related lengthening and boundary tones, and prominence, especially lexical stress (Katsika, 2016; Katsika, 2014; Kim, Jang, & Cho, 2017). Finally, it is unclear whether internal phonological properties of the word, such as the position of stress, plays a role in the kinematic manifestation of prominence. We begin to probe these issues here by examining the supralaryngeal kinematic profile of stress separately from accent as a function of position of the word in the phrase and position of the stressed syllable in the word in Greek. Stress is separated from accent by the means of words placed in unaccented, and specifically de-accented, positions, vs. accented ones. Special care is taken so that the accented positions do not bear contrastive focus in order to avoid extremely high degrees of prominence. Instead, broad and narrow focus positions are used, helping us also test the hypothesis that prominence phonetically encodes focus structure and not accentual status (cf. Mücke & Grice, 2014; Roessig & Mücke, 2019). In sum, three stress positions (antepenultimate, penultimate, ultimate; i.e., all possible stress positions in Greek), two accentual conditions (accented and de-accented) and two phrasal positions (phrase-medial and phrase-final) are considered. Furthermore, the accentual conditions come from two types of focus (broad and

narrow), while the de-accented conditions are by default unfocused. Our ultimate goal is to contribute to the field's understanding of the structural and phonetic nature of prominence and its relations to other prosodic dimensions, such as prosodic boundaries, while offering articulatory data of a less studied language.

1.2. Introducing key terminology

The development of metrical phonology (Lieberman & Prince, 1977; Nespor & Vogel, 1986; Pierrehumbert & Beckman, 1988; Selkirk, 1984) shaped the field's view of stress as a structural element of language. Stress is no longer a dimension that solely contrasts lexical items, but is one of the main devices that language possesses to encode hierarchical organization. In this view, prosody is a hierarchical structure, constructed by the means of grouping and prominence (e.g., Beckman & Pierrehumbert, 1986; Hayes, 1989; Nespor & Vogel, 1986; Selkirk, 1984). Grouping marks the edges of constituents of different sizes, such as words and phrases, organizing the string of speech into chunks of adequate size and type for cognitive processing. Prominence marks the heads of these constituents, meaning syllables within words (i.e., stressed syllables) and words within phrases (i.e., accented words) that are rhythmically or conceptually important, also facilitating language processing. Prosodic structure thus encodes hierarchies of grouping and prominence that are fundamental in the production, comprehension, acquisition and learning of language.

Numerous proposals have been put forward for the hierarchy of grouping, which, although not in agreement about the number of levels constituting the hierarchy, converge in postulating at least a minor and a major phrase level above the word level (see Shattuck-Hufnagel & Turk, 1996 for an overview). We refer to minor phrases as *intermediate phrases* (ip), and to major phrases as *Intonational Phrases* (IP), adopting the terms coined by the Autosegmental-Metrical model of intonational phonology (Beckman & Pierrehumbert, 1986; Pierrehumbert, 1980). The edges of these phrases are marked by specific pitch movements, among other types of markers. Intermediate phrases are marked by *phrase accents* and Intonational Phrases by *boundary tones*, using again the terms introduced by Autosegmental-Metrical Phonology. Since Intonational Phrases are higher in the hierarchy than intermediate phrases, the end of an Intonational Phrase always coincides with the end of an intermediate phrase, which means that a boundary tone is always preceded by a phrase accent. With respect to prominence, we refer to the sources of lexical and phrasal prominence as *stress* and *accent* respectively. Stress is marked for most languages at the prosodic word level, and accent is marked at the intermediate phrase level mainly via specific pitch movements, called *pitch accents*. The last pitch accent of a phrase is called *nuclear pitch accent*, with the ones preceding it been referred to as *pre-nuclear*.

1.3. The acoustic and articulatory correlates of prominence

Both grouping and prominence are associated with temporal and pitch modulations. In the temporal domain, speech units at phrase boundaries and under prominence involve longer acoustic durations and longer articulatory gestures than their phrase-medial counterparts (e.g., phrase boundaries: Beckman & Edwards, 1992; Byrd & Saltzman, 1998; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992; prominence: Cambier-Langeveld & Turk, 1999, Cho, 2006). In the tonal domain, pitch movements of different types mark accented

words and words that occupy positions adjacent to phrase boundaries (cf. Silverman et al., 1992).

Focusing on prominence and its tonal aspect, not all languages use fundamental frequency (F0) to mark lexical stress, and often it is not clear whether F0 is one of the correlates of lexical stress or of a higher level of prominence (cf. Fletcher, 2010 for an overview). On the other hand, there is general agreement that accented words are marked by pitch accents on the stressed syllable of the accented word (e.g., Beckman & Pierrehumbert, 1986; Silverman et al., 1992). Thus, accented syllables can be discretely distinguished from their unaccented counterparts on the basis of the presence of pitch accents. The degree of prominence on them might be related to the type of pitch accent, and some gradient aspects of pitch, such as pitch excursion (see Gussenhoven, 2004) and timing of the extrema (i.e., the low and high turning points) in pitch movements (cf. Ladd & Morton, 1997).

In the temporal domain, longer acoustic durations due to presence of prominence are detected in many languages (see overview in Fletcher, 2010). Relatedly to these durational effects, in some languages, unstressed vowels are reduced (e.g., Fourakis, 1991; Padgett & Tabain, 2005; Nowak, 2006). Languages differ in the extent they use duration to mark prominence, and it is unclear whether lengthening increases with degree of prominence. Most studies agree that accented syllables are longer than unaccented stressed syllables (e.g., Baltazani & Jun, 1999; Cambier-Langeveld & Turk, 1999; Cho & Keating, 2009; Cho & McQueen, 2005; de Jong, 2004; Turk & Sawusch, 1997; Turk & White, 1999), which are in turn longer than unaccented unstressed syllables (e.g., Arvaniti, 1992; Crystal & House, 1988; de Jong, 2004; de Jong & Zawaydeh, 2002; Heldner & Strangert, 2001; Rietveld, Kerkhoff, & Gussenhoven, 2004). There are also indications that units under contrastive focus lengthen more than units under broad focus (cf. Baumann, Grice, & Steindamm, 2006 for German). However, not all languages use duration equally and/or for both stress and accentuation (cf. de Jong & Zawaydeh, 2002; Dogil, 1999; Fant, Kruckenberg, & Nord, 1991; Oller, 1979). For example, Arabic shows longer stressed syllables than unstressed ones, but focus does not further affect duration of the stressed syllable (de Jong & Zawaydeh, 2002).

The acoustically detected prominence-induced lengthening is further supported by kinematic evidence. Studies on articulation have shown that prominence is correlated with longer, larger and faster articulatory movements than their non-prominent counterparts (Beckman & Edwards, 1994; Beckman et al., 1992; Cho, 2005, 2006; de Jong et al., 1993; de Jong, 1991, 1995; Fowler, 1995; Harrington et al., 1995, 2000). However, not all three kinematic dimensions, i.e., displacement (larger), duration (longer) and velocity (faster) consistently differentiate between accented and unaccented gestures in the literature. For instance, in Cho (2006) these patterns hold for the lip opening movement of the consonants in the test stimuli, which were all bilabial and in the onset position of the test syllables (cf. also de Jong, 1991, Fowler, 1995). The kinematic profile of the lip closing part of these consonants, on the other hand, depend on the context. In particular, when the accent is on the preceding vowel, the lips form a larger, longer and faster lip closure, whereas, when the accent is on the following vowel, the lip closure is larger, but not faster (except if the previous vowel is /a/) or longer (although the deceleration phase of lip closure was per se longer). Similarly,

in Beckman et al. (1992), accented syllables are produced with larger and longer, but not faster, jaw opening gestures as compared to their unaccented counterparts. Moreover, the property of faster articulatory movements has been attributed in the literature to either high velocity or shorter time-to-peak velocity (cf. Cho, 2006). Time-to-peak velocity has been used as a measure of the abstract dimension of stiffness (Byrd & Saltzman, 1998; used in e.g., Cho, 2002, 2006; Mücke & Grice, 2014). The shorter the time-to-peak velocity is, the higher the stiffness, and the faster the movement. Another empirical estimate of kinematic stiffness is the ratio of peak velocity to displacement, which captures the observation that peak velocity varies with displacement (Munhall, Ostry, & Parush, 1985; Ostry & Munhall, 1985; used in e.g., Beckman et al., 1992; Hawkins, 1992; Roon, Gafos, Hoole, & Zeroual, 2007). Despite these differences in the findings reported in the literature, it can be concluded that articulatory movements under prominence undergo spatiotemporal expansion, often referred to as strengthening (cf. Cho, 2006). Prosodic strengthening has also been associated with increased coarticulatory resistance, meaning strengthened movements are less prone to coarticulatory effects exerted by their neighboring articulations (e.g., Cho, 2004; de Jong et al., 1993). Support for extreme or augmented articulatory movements in prominent syllables comes predominantly from English, and is mainly found at high prosodic levels. Importantly, as highlighted by Mücke and Grice (2014), accent is not thoroughly disentangled from focus structure, and, as a result, the described effects might confound these two factors. Targeted examination of the possible combinations of accent and focus (absence of phrasal prominence, broad focus, narrow focus and contrastive focus) in German and English indicated that an increase in kinematic parameters (displacement, duration and velocity) is controlled by both focus structure and accentuation (German: Mücke & Grice, 2014; Roessig & Mücke, 2019; see also Hermes et al., 2008; English: Katsika et al., 2020). This increase occurs incrementally from unaccented words to broad focused, then narrowly focused, and finally contrastively focused ones (Roessig & Mücke, 2019; see also Katsika et al., 2019 for English). With respect to lower prosodic levels, there is evidence that strengthening also emerges from lexical stress, meaning when comparing full/stressed vowels to reduced/unstressed ones (e.g., Beckman & Edwards, 1994; Kelso, Vatikiotis-Bateson, Saltzman, & Kay, 1985).

At last, intensity, vowel quality, spectral tilt and formant patterns have also been detected as cues to prominence (e.g., Astruc & Prieto, 2006; Crosswhite, 2003; Nowak, 2006). Although these properties are beyond the scope of the current study, it is important to note that there is significant cross-linguistic variation in how languages phonetically express prominence. Languages might use a combination of all or some of the prominence-marking dimensions summarized here, with the extent of contribution of each dimension to signaling stress, accent (or focus) being also language-specific.

1.4. Accounting for prosodic strengthening

Prominence-induced strengthening has been attributed either to *sonority expansion* (Beckman et al., 1992) or *hyperarticulation* (de Jong, 1995). According to the sonority expansion account, accented vowels are produced by lowering the jaw over a sustained period of time, which enlarges the vocal tract and makes the vowel sound louder. As a result, the *sonority* of the vowel is *expanded*, and the percept of prominence is

enhanced. Alternatively, prominence-induced strengthening might be caused by localized hyperarticulation, i.e., by more extreme articulatory movements. For example, when accented, low vowels become lower (e.g., Cho, 2005; de Jong, 1995; Harrington et al., 2000). Similarly, consonantal productions under accent involve more extreme and extensive constrictions (e.g., Bombien, Mooshammer, Hoole, Rathcke, & Kuehnert, 2007; Cho & Keating, 2009; de Jong, 1995). There are indications of speaker-specific flexibility as to what dimension of a segment is produced with more extreme articulation. For instance, an accented /i/ might be produced either more fronted or higher than its unaccented counterpart (cf. de Jong, 1995; Harrington et al., 2000). Sonority expansion and hyperarticulation are not necessarily mutually exclusive, and could potentially be historically connected to each other in some languages/dialects (Harrington et al., 2000; see also Cho, 2005; de Jong et al., 1993; Harrington et al., 2000). However, the two theories could possibly make different predictions when considering prominence on high vowels, such as /i/. The hyperarticulation hypothesis predicts higher tongue body, while the sonority expansion hypothesis lower tongue body by virtue of lowering the jaw (although note that, albeit effortful and somewhat challenging, it is technically and/or psychologically possible for the tongue to be positioned high while the jaw is low).

Several proposals have been put forward for the articulatory control of prominence. Harrington et al. (1995) tested two alternative models of the differences between unaccented and accented (bearing contrastive focus) vowels: (a) *truncation*: the relative timing between the closing and opening jaw gestures of unaccented vowels is such that their temporal overlap increases, and the duration of the vowel decreases (cf. Edwards, Beckman, & Fletcher, 1991; Beckman et al., 1992; de Jong et al., 1993), and (b) *rescaling*: proportional decrease in displacement and duration resulting in shrinking of the unaccented vowels as compared to the accented ones. Their findings showed that truncation corresponds better to human productions. Cho (2006) further evaluated truncation and rescaling, and added *target modification* and *stiffness* as the four possible parameters controlled for prosody-driven modifications. In target modification, articulatory gestures with more extreme targets would have increased displacement and peak velocity increasing in proportion to displacement, and as a result the duration to the modified target would not change. Setting a gesture's stiffness to a lower value, on the other hand, would result in slowing down the movement, and thus to a longer time to their target, which should, in turn, remain unmodified. None of these four parameters (truncation, rescaling, target modification and stiffness) was identified as the sole driver of the prominence-related effects, whereas every single one of them could be contributing in some sort of combination to these effects. This lack of a single parameter for inducing prominence is further enhanced by evidence that the prominence-related kinematic effects, and thus the prominence-inducing strategies as well, are speaker-dependent (Cho, 2006, but also see Avesani, Vayra, & Zmarich, 2007; Beckman et al., 1992; Cho, 2005, de Jong, 1995, Dohen & Loevenbruck, 2005; Dohen, Loevenbruck, & Hill, 2006; Harrington et al., 1995, 2000; Hermes et al., 2008; Mücke & Grice, 2014).

Recent advances in Articulatory Phonology (e.g., Browman & Goldstein, 1986, 1992) and the Task Dynamics model of sensorimotor control and coordination (Saltzman & Munhall, 1989) have proposed the concept of μ -gestures (Saltzman, Nam, Krivokapi , & Goldstein, 2008) as the vehicle of controlling the production of prominence, expanding

on the p-gesture model of prosodic boundaries (Byrd & Saltzman, 2003). In Articulatory Phonology, phonological representations of segments are composed of linguistically relevant constricting events, called *gestures*, that control the speech organs. There are three gesture types: constriction gestures, tone gestures and modulation gestures. Constriction gestures are specified for abstract linguistic tasks (such as labial closure), are realized by coordinated actions of specific articulators (i.e., the jaw, lips, tongue, velum and glottis), have spatio-temporal properties, and are triggered by internal oscillators that are coupled to each other either in-phase (synchronously) or anti-phase (sequentially) (e.g., Goldstein, Byrd, & Saltzman, 2006). Tone gestures are similar to constriction gestures: they are specified for F0 targets, are realized by the coordinated actions of articulators, such as the lungs, trachea, larynx and various muscles and are coordinated with other gestures. μ -gestures, standing for modulation gestures, are different from either constriction or tone gestures in that they are not related to specific articulators and do not have specific constriction task/target. Their role is to alter the spatiotemporal properties of the constriction gestures that are concurrently active with them (Saltzman et al., 2008). In particular, μ -gestures are either temporal, modulating the rate of utterance time flow, much like their predecessors π -gestures, or spatial, augmenting or decreasing the movements of co-active constriction gestures by smoothly changing their target parameter. This model has been used to date for modeling lexical stress and polysyllabic shortening (Saltzman et al., 2008). It is yet to be determined how exactly the activation of these gestures is different between stress and accentuation, or among different levels of prominence, or between prominence and boundaries.

1.5. Prominence in Greek

This section briefly describes prominence in Greek (for an overview see Arvaniti, 2007). Greek lexical stress is contrastive and placed in one of the three final syllables of a word. Its exact position is phonologically unpredictable, and, although somewhat constrained by morphological criteria, it remains an acquired lexical dimension. A stress minimal pair and a stress minimal triplet are offered in list (A) as examples.

- (A) 1. [ˈnɔmɔs] “law, n.” – [nɔˈmɔs] “county, n”
 2. [tiˈlɛfɔnɛ] “telephones, n.” – [tilɛˈfɔnɛ] “telephone, 2nd prs. imp.” – [tilɛfɔˈnɛ]
 “telephones,
 3rd prs. ind.”

In terms of phonetic correlates, stressed vowels in Greek have longer acoustic duration and higher acoustic amplitude than their unstressed counterparts (e.g. Arvaniti, 1991, 2000; Baltazani, 2007; Botinis, 1982, 1989; Dauer, 1980; Fourakis, Botinis, & Katsaiti, 1999; Nicolaidis & Rispoli, 2005). In terms of duration, the whole stressed syllable is affected, although the effect on the consonant is rather inconsistent (e.g., Arvaniti, 2000; Botinis, Fourakis, Panagiotopoulou, & Pouli, 2001a; Botinis, Fourakis, & Bannert, 2001b). There is evidence for a cue trading relationship between acoustic duration and acoustic amplitude in signaling Greek stress (Arvaniti, 1991, 2000; Dauer, 1980), which means that stressed vowels in Greek might be one of the following: (1) longer, (2) louder or (3) both longer and louder than their unstressed counterparts. Moreover, vowel space expands when vowels

are stressed as opposed to unstressed, indicating that stressed vowels are hyperarticulated (Fourakis et al., 1999; Nicolaidis & Rispoli, 2005). In parallel, there is evidence that unstressed vowels are centralized (e.g., Baltazani, 2007; Fourakis et al., 1999; Nicolaidis, 2003; but this is not supported in Arvaniti, 1991, 2000; Dauer, 1980). If unstressed vowels are indeed centralized, this is presumably due to undershoot (cf. discussion in Arvaniti, 2007).

Fundamental frequency (F0) does not mark stressed syllables directly, but indirectly, in the sense that it is the stressed syllable of the accented word that bears the pitch accent (e.g., Arvaniti, 2000; Botinis, 1998). According to the Greek Tone and Break Indices system (GrToBI: Arvaniti & Baltazani, 2005), there are four types of nuclear pitch accents in Greek: L*, H*, L + H* and H*+L. The F0 peaks of the H* and L + H* and the F0 minimum of L* co-occur with the stressed vowel, while the F0 peak of the H*+L occurs just before the stressed syllable (Arvaniti & Baltazani, 2005; Arvaniti, Ladd, & Mennen, 2006).

In addition to the presence of pitch accent, focused words and/or syllables seem to be acoustically longer than their non-focused counterparts (the whole focused word: Baltazani & Jun, 1999; stressed syllable of focused words, but not the whole word: Botinis, Bannert, & Tzimokas, 2002; Botinis & Bannert, 2003; Kastrinaki, 2003). However, some results do not show focus-related lengthening (Botinis et al., 2001b, 1995, 2001a).

1.6. Questions and hypotheses

The current study addresses the following question: What are the articulatory correlates of prominence in Greek, separating lexical stress from phrasal accentuation, as a factor of stress position in the word and word position in the phrase? Phrasal accentuation is examined by contrasting accented to de-accented conditions. Two types of focus are used to elicit accent, meaning broad and narrow focus, addressing the additional question of whether the phonetic profile of prominence encodes focus structure as well. The specific hypotheses tested are:

Hypothesis I: General kinematic profile of prominence. We expect that gestures under prominence in Greek will be longer, larger and faster than their non-prominent counterparts, extending previous findings on other languages (English: Beckman et al., 1992; Beckman & Edwards, 1994; Cho, 2005, 2006; de Jong, 1991, 1995; de Jong et al., 1993; Fowler, 1995; Harrington et al., 2000, 1995; German: Hermes et al., 2008; Mücke & Grice, 2014; Italian: Avesani et al., 2007). The patterns on the time dimension are further supported by evidence for longer acoustic durations under prominence in a wide range of languages (e.g., de Jong & Zawaydeh, 2002; Gussenhoven, 2009; Kohler, 1983; Turk & White, 1999; see also support for reduced unstressed vowels in e.g., Fourakis, 1991; Padgett & Tabain, 2005; Nowak, 2006, including Greek, e.g., Arvaniti, 1991). However, not all three dimensions (i.e., position, duration, and velocity) might be affected by prominence (cf. Beckman et al., 1992; Cho, 2006). In fact, the dimensions of position and duration might be in a cue trading relationship, based on discussion of Greek acoustic data in Arvaniti (1991, 2000) and Dauer (1980).

Hypothesis II: Hierarchy of prominence. We further hypothesize that kinematics will reflect a hierarchy of prominence. Specifically, we expect that stressed syllables will be longer than unstressed syllables (e.g., Beckman & Edwards, 1994; Crystal & House, 1988; de Jong, 2004; Kelso et al., 1985), and that accented syllables will in turn be longer than unaccented ones (e.g., Baltazani & Jun, 1999; Cambier-Langeveld & Turk, 1999; Cho & Keating, 2009; Cho & McQueen, 2005; de Jong, 2004; Turk & Sawusch, 1997; Turk & White, 1999; but see de Jong & Zawaydeh, 2002). In Greek specifically, the first part of this hypothesis concerning the distinction between stressed and unstressed syllables is directly supported by acoustic evidence (e.g., duration: Arvaniti, 1991, 2000; Baltazani, 2007; Botinis, 1982, 1989; Dauer, 1980; Fourakis et al., 1999; Nicolaidis & Rispoli, 2005). Moreover, there is evidence that Greek vowels are hyperarticulated when stressed (Fourakis et al., 1999; Nicolaidis & Rispoli, 2005), and centralized when unstressed (e.g., Baltazani, 2007; Fourakis et al., 1999; Nicolaidis, 2003; but this is not supported in Arvaniti, 1991, 2000; Dauer, 1980; see also discussion in Arvaniti, 2007), suggesting larger gestures under stress. With respect to the second part of this hypothesis that addresses the distinction between accented and unaccented syllables, the discussion based on Greek data is inconclusive. Some studies have detected accentual lengthening (e.g., Baltazani & Jun, 1999; Botinis et al., 2002; Botinis & Bannert, 2003; Kastrinaki, 2003), while others have not (Botinis et al., 2001a, 2001b, 1995). These inconsistencies are presumably due to the different types of stimuli, including different types of focus, these studies used. The current study places the accented words either in broad or in narrow focus, but not in contrastive focus. In this way, extreme cases of prominence are avoided. Greek might be similar to German and English in how kinematics reflects focus structure rather than just status of accentuation (Baumann et al., 2006, Mücke & Grice, 2014; Roessig & Mückem, 2019; English: Katsika et al., 2020). In this case, we expect not only that accented gestures will be longer than unaccented ones, but more specifically that they will be longer in narrow focus as compared to broad focus. Extending Hypothesis II to the other kinematic dimensions, we expect modifications in displacement and velocity to increase cumulatively with prominence degree (cf. Baumann et al., 2006, Mücke & Grice, 2014; Roessig & Mückem, 2019; English: Katsika et al. 2020).

Hypothesis III: Position of stress in the word. It is unclear whether the articulatory correlates of stress and accent vary with position of the stressed syllable within the word. With Greek stress being a lexical property, it makes conceptual sense to expect stress to have an unaltered kinematic signature.

Hypothesis IV: Position of word in the phrase. As for the position of the word within the phrase (medial or final), interaction effects are predicted when a syllable or word is both prominent and phrase-final. This is because phrase-final positions are similar to prominent positions in that they involve longer gestures (e.g., Byrd, 2000; Byrd, Kaun, Narayanan, & Saltzman, 2000), albeit the phrase-final gestures are slower than their phrase-medial counterparts while accented gestures are expected to be faster than unaccented ones (Cho, 2006). In parallel, interactions between lexical stress and phrasal position have already been found with respect to the timing of boundary-related effects in Greek (Katsika, 2016; Katsika, 2014). Note that this work has further suggested that concurrence of phrase-final

and prominent positions does not result in the two lengthening effects (prominence- and boundary-related) functioning additively.

Table 1 below summarizes these hypotheses. Our goal is to discover the supralaryngeal articulatory signatures of lexical stress and phrasal accent in languages like Greek that use stress contrastively, and the effects of word and phrase boundaries on them. Ultimately, we aim to understand the architecture of prosodic structure and the type of spatio-temporal articulatory control that gives rise to it.

2. Methods

The speech data used here are the same as in Katsika (2014) and Katsika (2016). This section summarizes the shared methods among the three studies and highlights the methods that are specific to the current study.

2.1. Participants

Eight native speakers of standard Greek (5 female, 3 male) participated in this set of experiments. At the time of their participation, they were all associates of Yale University, ages between 19 and 31, who had been in the United States of America between 1 and 6 years. Participants were naïve to the purpose of the experiment, and reported no speech, hearing or vision problems. Participants gave informed consent, approved by the Yale University Human Investigation Committee, and received financial compensation for their participation. Based on a set of post-processing analyses, data from three of the participants were judged inappropriate for further analysis (details to be presented below). We thus report here the results of the five remaining participants.

2.2. Design and stimuli

To examine the effect of lexical stress and its position within the word on articulatory gestures the following stress minimal triplet of made-up words was used: *MAmima* ('mɛmimɛ), *maMIma* (mɛ'mimɛ) and *mamiMA* (mɛmi'mɛ) (capital letters stand for lexical stress). These words were constructed to have identical segments, but different stress position and different meaning. *MAmima* is stressed on the antepenult (S1: stress-initial word), *maMIma* is stressed on the penult (S2: stress-medial word), and *mamiMA* is stressed on the ultima (S3: stress-final word). In terms of meaning, each of these words is a different narcotic plant. These meanings were selected in order for the words to make sense in all of their frame sentences.

The effect of accent was separated from the effect of stress by the means of two sets of frame sentences. In one set of frame sentences, the test words *MAmima*, *maMIma* and *mamiMA* were accented (A) bearing the nuclear pitch accent, while in the other set, the test words were de-accented (D) following the nuclear pitch accent by several words. The sentences creating the accented condition were of the following types: affirmative declaratives (AD), yes–no questions (YNQ), causative clauses (CC) and parenthetical clauses (PC) (see Table 2, rows 5–8 for respective examples). In these sentences, the accented words were not contrastively focused. Instead, the accented words conveyed narrow focus (NF) in affirmative declaratives (AD) and parenthetical clauses (PC), and

broad focus (BF) in yes–no questions (YNQ) and causative clauses (CC). The set of sentences satisfying the de-accented condition were negative declaratives showing reservation (ND), wh-questions (WhQ) and imperative requests (IR) (see Table 2, rows 1–3 for respective examples). In the de-accented conditions, the test words were by default unfocused (UN). Employing a large number of sentence types served to increase variability in the data, but primarily to elicit a representative set of boundary tones in order to assess boundary tone coordination, reported in Katsika (2014).

The frame sentences placed the test words in phrase-final position (IP). To detect any interaction of prominence effects with phrasal position, two additional sets of stimulus sentences were constructed in which the test words were in phrase-medial positions (W). A set of affirmative declaratives placed the test words in phrase-medial accented positions bearing narrow focus (see example in Table 2, row 9), and a set of negative declaratives showing reservation placed them in phrase-medial de-accented positions (see example in Table 2, row 4). There was no counter frame sentence for the broad-focused phrase-final condition, since, by definition, a word cannot bear the nuclear pitch accent of broad focus while being phrase-medial.

Some other controlled features of the experimental design were the following: The words before and after the test word were the same across all frame sentences. The preceding word was /'ɛkəpi/ ('raw') and the word following was /metə'ks i/ ('among'), positioning neighboring lexical stresses as far as possible from the test word (i.e., two syllables away on each side). In addition, all frame sentences had seven syllables preceding the test word and thirteen syllables following it. In the cases in which the test word was phrase-final (IP), the thirteen syllables following it constituted a separate phrase. To elicit the targeted intonation contour, contextualizing sentences (see Table 2, sentences in italics) preceded all stimulus sentences (shown in IPA in Table 2) except yes–no questions and causative clauses. The two latter types did not need context in order to be uttered with the targeted intonation.

For each of the 3 test words, 9 frame sentences were employed, yielding 27 stimulus sentences. The experimental protocol contained 9 blocks, each including the 27 stimulus sentences in a randomized order, yielding 243 test utterances in total. Table 1 (adapted from Katsika, 2016) lists the 9 frame sentences. For demonstration purposes, the examples shown use the test word *MAmima*.

2.3. Apparatus and recording procedure

The kinematic recordings took place at the physiology lab at Haskins Laboratories, using the AG500 three-dimensional electromagnetic transduction device (Carstens Medizinelektronik). The movement of the tongue dorsum, tongue body, tongue tip, upper lip and lower lip were tracked by the means of receiver coils. For the purposes of post-processing (e.g., head correction), receiver coils were also attached to flesh points representing the following organs: the upper incisor, left side of the jaw, front side of the jaw, left ear, right ear, and nose. Following Hoole, Zierdt, and Geng (2003), each experimental session was preceded by the standard calibration procedure. A Sennheiser shotgun microphone was used to acquire acoustic data in parallel with the kinematic

recordings. The micro-phone was set at a sampling rate of 16 kHz, and was positioned roughly 12 inches away from the participant.

A short training session (20–30 min) preceded the experimental session by 1–3 days. The purpose of the training session was to familiarize the participants with the made-up words, the targeted intonational contours, and the stimuli presentation. For the experimental session (2–3 h), custom software (developed by Mark Tiede, Haskins Laboratories) was used to present the instructions and the speech materials on a computer screen positioned roughly 60 inches away from the participant. The instructions directed the participants to pay attention to the position of lexical stress on the test words – Greek orthography has a specific symbol (‘ ’) for stress, e.g., ‘*μάμια*’, ‘*μαμίμα*’, ‘*μαμίμά*’ – the punctuation signs in order to achieve the correct phrasing, and the words in bold font, which were “the words bearing the main information of the sentence”, meaning the words under focus. For each trial, the contextualizing sentence appeared first and in green font, and the test sentence appeared in blue font 10 seconds later. The participants were instructed to read the contextualizing sentence silently and the test sentence aloud. They were also instructed to speak at their normal speech rate. Sentences produced with speech errors, interruptions or disfluencies were repeated.

2.4. Analyses

2.4.1. Pre-processing analyses—The data were tested for reliability, smoothed, corrected and translated to the occlusal plane using the TAPADM pre-processing procedure (Three-dimensional Articulographic Position and Align Determination with MATLAB™, developed by Andreas Zierdt; cf. Hoole et al., 2003). An additional preliminary analysis determined whether the data were produced with the targeted F0 contour by the means of acoustic inspection. Based on these two sets of analysis, the data of three speakers were considered inadequate for further analysis: the data from one speaker included a large amount of dropped or noisy data from at least one receiver coil, and the data from the other two speakers were not produced with the targeted F0 contours. Specifically, the latter two speakers instead of producing the de-accented conditions with the targeted L-H% combination of phrase accent and boundary tone, they used an alternative contour for these constructions (L–L%) (cf. Arvaniti & Baltazani, 2005; see also Arvaniti & Ladd, 2009 for wh-questions; Baltazani, 2006 for negative declaratives).

The data from the remaining five participants, referred to here as F01, F02, F03, F04 and M05 (four female and one male), were subject to further kinematic and intonational pre-processing analyses. Based on the kinematic analysis, less than 3% of the data were disregarded due to abnormalities in their displacement or velocity signal. The intonational analysis examined the data for their prosodic boundaries using GrToBI (Arvaniti & Baltazani, 2005). On the basis of this analysis some tokens were excluded due to missing or wrong boundary production. As a result, there were between 5 and 15 tokens per test word in each frame sentence per speaker, with the exception of F04’ parenthetical clauses, which were all excluded due to missing prosodic boundaries. Some frame sentences had more than 9 repetitions, which was the number of blocks in the experimental protocol. This was because additional repetitions were acquired over the course of the experiment following

resumption of the recording after interruption due to software error or after the participant's request for a break.

To confirm accurate characterization of intonational contours into accented and de-accented, a second labeler performed a final pass of the resulted dataset, which consisted of two steps. First, the second labeler used GrToBI to annotate prosodic boundaries and detect pitch accents. Their labels were in complete (100%) agreement with the original analysis. Second, the labeler marked the acoustic onset of each test utterance at the moment in which phonation is initiated and the acoustic offset of the test word at the end of F2 of the final vowel of the test word. On the basis of these common events, the Lucero, Munhall, Gracco, and Ramsay (1997) nonlinear time warping technique was used to compute the normalized alignment of F0 signals of the acquired utterances per accent type (accented vs. de-accented). This analysis included only the phrase-final conditions in order for phrasal tones to coincide with respect to their position in the phrase. The resulted mean F0 signals are shown in Fig. 1, which confirms that the test utterances were accurately categorized into accented and de-accented. Specifically, the utterances marked as de-accented involve a low F0 stretch from the nuclear pitch accent that begins the utterance to the boundary tone (!H%) that marks the end of the utterance. The test word is phrase-final, occurring during the low F0 stretch and carrying the boundary tone. Note that the average of the accented conditions involved constructions with different types and locations (due to the various stress positions) of pitch accents, while all de-accented conditions targeted the same contour (L*H L-!H%) and position of phrasal tones.

2.4.2. Kinematic analyses—The filtered data were semi-automatically labeled for kinematic landmarks, using custom software (Mark Tiede, Haskins Laboratories). Specifically, the C and V gestures of the words *MAmima*, *maMIma* and *mamiMA* were marked on lip aperture (Euclidian distance between upper and lower lip) and tongue dorsum vertical displacement (with respect to the bite plane) respectively for the following kinematic timepoints (see Fig. 2): onset, peak velocity, target, constriction maximum and release. Timepoints were identified on the basis of velocity criteria, i.e., peak velocity for the homonymous timepoints, velocity minima for constriction maxima, and velocity plateaus for the other timepoints. Velocity plateaus were detected based on a set threshold of velocity range between two consecutive alternating velocity extrema (i.e., one minimum and one maximum). This threshold was 20% for all timepoints, except for C onset and C offset, for which the threshold was 10%, due to the small amplitude of lip aperture for the consecutive labial C constrictions (Fig. 3).

Using these timepoints, the measures listed in (B) were calculated.

- (B) 1. Formation duration: the temporal interval (in ms) between onset and release, following the split-gesture hypothesis (Nam, 2007).
2. Displacement: defined as maximal displacement, and calculated as the spatial difference (in mm) between the point of maximal constriction and gestural onset.
3. Peak velocity: Formation's peak velocity (in cm/sec).

4. Ratio between peak velocity and displacement as an estimate of stiffness (in (cm/sec)/cm, i.e., Hz) (Munhall et al., 1985; Ostry & Munhall, 1985).

2.4.3 Statistical analyses—Two data frames were constructed, one for C and one for V gestures. Within each data frame, linear mixed effects models were fitted to each dependent variable (formation duration, peak velocity, and displacement) using the `lmerTest` package in R (Kuznetsova, Brockhoff, & Christensen, 2017; R Core Team, 2019) with a Satterthwaite approximation of the degrees of freedom. For each dependent variable, the model selection process started with the maximal structure of fixed and random effects, and concluded with the minimal adequate model by the means of the `drop1` function. The fixed factors considered are presented in list (C).

- (C) 1. Stress position, with levels S1 (stress-initial), S2 (stress-medial), S3 (stress-final)
2. Focus type, with levels unfocused (UF), broad focus (BF), and narrow focus (NF)
3. Boundary type, with levels word (W), and Intonational Phrase (IP)
4. Syllable affiliation of target C or V gesture, with levels C1 (C gesture of initial syllable), C2 (C gesture of medial syllable) and C3 (C gesture of final syllable) for C gestures, and V1 (V gesture of initial syllable), V2 (V gesture of medial syllable) and V3 (V gesture of final syllable) for V gestures.

Presence of stress, accentual status, vowel quality, and construction type were not entered as factors, because they either overlap significantly with one of the factors listed above or with some combination of them. For example, presence of stress is captured by the combination of factors stress position and syllable affiliation. A gesture in S1 stress position and initial syllable affiliation is stressed word-initially and contrasts directly with gestures in S2 and S3 stress positions with initial syllable affiliation, which are unstressed in word-initial positions. Using the combination of stress position and syllable affiliation for capturing stress effects has the additional advantage of informing us about the scope (stretch of speech affected) by the stress effects, since it can reveal finer distinctions between the two unstressed conditions. Presumably, unstressed gestures that are adjacent to stressed ones could be undergoing anticipatory or spill-over stress-induced effects, depending on whether they precede or follow the stress (cf. Dimitrova & Turk, 2010). Such effects should attenuate or even disappear on more remote gestures. To further illustrate the selection of factors to consider, vowel quality, /i/ is uniquely represented by the level “medial” of the syllable affiliation factor (C2, V2), and unaccented gestures completely overlap with the level “unfocused” of the focus factor.

In order to fully assess effects of stress both paradigmatic and syntagmatic pairwise comparisons were assessed. In paradigmatic comparisons, gestures in a given position within the word were compared across word types. For example, word-initial C gestures were compared across stress-initial, stress-medial and stress-final words). The syntagmatic approach compared a given C or V gesture to the other C or V gestures respectively of the

same word. For instance, word-initial C gestures were compared to their word-medial and word-final counterparts.

To control for the expected covariance between duration and stiffness, the models of duration included in addition to the categorical factors in list (C), the ratio of peak velocity over displacement as a numeric predictor. As a reminder, this ratio has been used as an estimate of stiffness and has been shown to increase as movement duration decreases (Munhall et al., 1985; Ostry & Munhall, 1985). Similarly, in order to account for the expected covariance between peak velocity and displacement, the model of displacement included peak velocity as a numeric predictor, and vice versa, the model of peak velocity included displacement as a numeric predictor. Pairwise comparisons were conducted via the emmeans package using the Holm adjustment for multiple comparisons (α level = 0.05).

3. Results

The results for each kinematic parameter (duration, displacement and velocity) are presented in Sections 3.1, 3.2 and 3.3 respectively. In each of these sections, the minimal adequate model of the respective dependent variable is listed first, followed by a description of its relationship with the numeric predictor. Next, the effects of each fixed factor are described in the following order: stress, accent/, and boundaries.

3.1. Duration

3.1.1. Formation duration of C gestures—The minimal adequate model of formation duration for C gestures included random intercepts and random slopes for syllable affiliation per speaker. The significant fixed effects of the model are summarized in Table 3.

3.1.1.1. Relationship between duration and stiffness.: This model confirmed that, everything else kept equal, duration of C gesture formation decreased linearly with stiffness, estimated by the ratio of peak velocity over displacement, with a slope of 2.26 ms/Hz (see Fig. 4 (a); Fig. 4 (b) is addressed in Section 3.1.2.1).

3.1.1.2. Effects of stress.: The pairwise comparisons testing interaction (7) in Table 3 revealed that, when keeping stiffness stable at its mean value ($\text{mean}_{\text{stiffness}} = 26.1$ Hz), stressed C gestures were longer than their unstressed counterparts. As illustrated in Fig. 5, C gestures of initial syllables were longer in stress-initial words than in either stress-medial or stress-final words (C1: $S1 > S2$ with $\beta = 14.3$, $SE = 1.0$, $p < 0.05$, and $S1 > S3$ with $\beta = 17.5$, $SE = 1.0$, $p < 0.05$). Similarly, C gestures of medial syllables were longer in stress-medial words than in either stress-initial or stress-final words (C2: $S2 > S1$ with $\beta = 7.7$, $SE = 1.2$, $p < 0.05$, and $S2 > S3$ with $\beta = 16.8$, $SE = 1.1$, $p < 0.05$). Lastly, C gestures of final syllables were longer in stress-final words than either stress-initial or stress-medial words (C3: $S3 > S1$ with $\beta = 17.9$, $SE = 1.2$, $p < 0.05$, and $S3 > S2$ with $\beta = 4.6$, $SE = 0.9$, $p < 0.05$). The pairwise comparisons testing interaction (4) confirmed that stressed syllables were longer than their unstressed counterparts regardless of the position of stress within the word and the position of the word in the phrase ($p < 0.05$ for all comparisons). The only exception was the case of word-final C gestures (C3) of phrase-final (IP) words, which were not significantly longer when stressed (S3) as opposed to unstressed with the stress on the

immediately preceding syllable (S2), presumably due to a combination of spillover effects exerted by the penultimate stress (S2) with phrase-final lengthening (see Fig. 6).

Crucially, as Fig. 5 illustrates, C gestures belonging in the two unstressed syllables were also different from each other, with those belonging to syllables adjacent to the stressed one being longer than those belonging to more remote syllables (see also Katsika & Tsai, 2019). In particular, C gestures of initial syllables were longer in stress-medial words (i.e., when stress immediately followed the target consonant) than in stress-final words (i.e., when stress was one syllable away from the target consonant) (C1: S2 > S3 with $\beta = 3.2$, $SE = 0.9$, $p < 0.05$), and C gestures of final syllables were longer in stress-medial words (i.e., when stress immediately preceded the target consonant) than in stress-initial words (i.e., when stress was one syllable away from the target consonant) (C3: S2 > S1 with $\beta = 13.4$, $SE = 1.2$, $p < 0.05$). C gestures of medial syllables were longer in stress-initial words as compared to stress-final words, suggesting that spillover effects are stronger than anticipatory effects (C2: S1 > S3 with $\beta = 9.1$, $SE = 1.4$, $p < 0.05$).

The patterns observed in stress-adjacent syllables suggest that the effects of stress extend beyond the stressed syllable to at least the preceding and the following syllable, with spillover effects being larger than anticipatory ones. These finer effects of stress proximity hold in both phrase-medial (Word) and phrase-final (IP) positions, as Fig. 6 demonstrates (with $p < 0.05$ for all pairwise comparisons).

As a point of clarification, we use here the terms *anticipatory* and *spillover* descriptively, i.e., to denote effects before and after the stressed syllable respectively. Such effects are clearly apparent in the paradigmatic comparisons described above (i.e., when gestures in a given position within the word were compared across word types). The syntagmatic comparisons (which examine gestures with respect to the other gestures of the same word) showed that stress-initial and stress-final words presented the expected pattern of the stressed C gesture of the word being longer than its unstressed ones (S1: C1 > C2: $\beta = 19.1$, $SE = 2.2$, $p < 0.05$, C1 > C3: $\beta = 22.8$, $SE = 2.8$, $p < 0.05$; S3: S3: C3 > C1: $\beta = 12.5$, $SE = 2.7$, $p < 0.05$, C3 > C2: $\beta = 23.3$, $SE = 1.4$, $p < 0.05$). However, neither stress-initial words nor stress-final ones demonstrated stress proximity effects. In fact, in stress-final words it was initial C gestures that were longer than stress-adjacent, medial ones (S3: C1 > C2: $\beta = 10.8$, $SE = 2.1$, $p < 0.05$), pointing to effects of secondary stress (cf. Arvaniti, 1992). Interestingly, stress-medial words did not present any distinctions of stress, with all of their C gestures being of equal length to each other. Given that all C gestures considered were labial stops (/m/), these patterns cannot be attributed to inherent segment durations, although it needs to be noted that the vowel context of these C gestures differed (/i/ for the medial C gesture, and /a/ for the initial and final ones).

Examining phrase-final and phrase-medial positions separately, according to interaction 4 in Table 3, confirmed these conclusions (see Fig. 6). In stress-initial and stress-final words, C gestures of stressed syllables were the longest ($p < 0.05$ for all pairwise comparisons), regardless of phrasal position. Stress-final words presented additional lengthening on their initial C gesture, induced presumably from secondary stress, both phrase-medially and phrase-finally ($p < 0.05$ for all pairwise comparisons). Stress-medial words did not present

any stress-related durational effects. They did nonetheless show, as expected, phrase-final lengthening on their final C gesture, as did the final C gestures of stress-initial and stress-final words ($p < 0.05$ for all pairwise comparisons). A new pattern that arose from investigating boundary type was spillover effects in stress-initial words, since their medial C gestures were longer than their final ones in phrase-medial positions (S1 in W: C2 > C3, with $\beta = 11.7$, $SE = 2.7$, $p < 0.05$). We take this to mean secondary stress is required in trisyllabic prosodic words with final primary stress, being also in accordance with a word demarcation account as per Turk and Dimitrova (2012).

In sum, according to both paradigmatic and syntagmatic comparisons, stressed C gestures are longer than their unstressed counterparts, supporting Hypothesis IIa. Paradigmatic comparisons further suggest both anticipatory and spillover effects of stress, while syntagmatic comparisons bring forward effects of secondary stress.

3.1.1.3. Effects of focus and accent.: Contrary to Hypothesis IIb and c, presence of accent and/or focus did not induce further prominence-related lengthening on the stressed C gesture (no interaction among the factors of focus type, stress position and syllable affiliation was detected). However, accent was found to cause some finer overall effects (Table 3, interaction 2), with C gestures of accented words under either broad or narrow focus being roughly 3.5 ms longer than in unaccented words (BF > UF, with $\beta = 4.6$, $SE = 0.5$, $p < 0.05$; NF > UF, with $\beta = 4.5$, $SE = 0.5$, $p < 0.01$).

Broad focus and narrow focus were not significantly different from each other, unless syllable affiliation was considered (Table 3, interaction 1), revealing some minimal differences between the two focused conditions, illustrated in Fig. 7. Medial C gestures were slightly longer under narrow focus as opposed to broad focus (C2: NF > BF, with $\beta = 2.1$, $SE = 0.8$, $p < 0.05$), while final C gestures presented the opposite pattern (C3: BF > NF, with $\beta = 2.4$, $SE = 0.8$, $p < 0.05$). Initial C gestures did not distinguish between broad and narrow focus. A look at accent/focus distinctions by examining their interaction with stress position and boundary type (Table 3, interaction 3) confirmed that C gestures were longer by an average of 4.7 ms in accented words (either with narrow or broad focus) as opposed to de-accented ones, regardless of the word's position in the phrase or which of its syllables was stressed ($p < 0.05$ for all pairwise comparisons). Nonetheless, these finer comparisons did not present any distinction between the two accented conditions (broad and narrow focus).

To summarize, our results show some minimal effects of phrasal prominence. Contrary to predictions (Hypothesis IIb and c), these effects do not increase the duration of the stressed C gesture. Instead, it is the temporal profile of the word overall that is affected. Importantly, in our data the duration of C gestures encode accentual status, and not focus structure as suggested by findings from German (Mücke & Grice, 2014; Roessig & Mücke, 2019; see also Hermes et al., 2008) and English (Katsika et al., 2020). These durational patterns of accent present the expected direction of accented gestures being longer than the unaccented ones, but they are small in magnitude (shorter than 5 ms).

3.1.1.4. Effects of phrase boundaries.: As predicted by Hypothesis IV, our analyses detected an interaction between position of lexical stress and phrase boundaries (Table 3, interaction 4), demonstrated in Fig. 8. According to our results, word-final C gestures underwent boundary-related lengthening unless the word-final syllable is stressed (C3 in S1: IP > W, with $\beta = 14.4$, $SE = 1.9$, $p < 0.05$; C3 in S2: IP > W, with $\beta = 10.4$, $SE = 1.3$, $p < 0.05$). In addition, stress-initial words had shorter initial C gestures when in phrase-final positions as opposed to phrase-medial ones (C1 in S1: W > IP, with $\beta = 8.4$, $SE = 1.4$, $p < 0.05$). These results confirmed the findings in Katsika (2016), which followed a different type of analysis of the same data.

3.1.2. Formation duration of V gestures—The minimal adequate model of formation duration for V gestures included random intercepts and random slopes for syllable affiliation per speaker, and the significant interactions of fixed factors summarized in Table 4.

3.1.2.1. Relationship between duration and stiffness.: This model detected a linear decrease of V gesture duration with stiffness with a slope of 3.1 ms/Hz, as shown in Fig. 4b.

3.1.2.2. Effects of stress.: The minimal adequate model included effects of lexical stress (Table 4, interaction 6). At mean value of stiffness ($\text{mean}_{\text{stiffness}} = 16.9$ Hz), V gestures of stressed syllables were longer than their unstressed counterparts. As illustrated in Fig. 9, word-initial V gestures were longer in stress-initial words (S1) than in either stress-medial (S2, with $\beta = 31.2$, $SE = 1.9$, $p < 0.05$) or stress-final (S3, with $\beta = 53.5$, $SE = 2$, $p < 0.05$) words, word-medial V gestures were longer in stress-medial words (S2) than in either stress-initial (S1, with $\beta = 30.3$, $SE = 1.7$, $p < 0.05$) or stress-final (S3, with $\beta = 25.3$, $SE = 1.7$, $p < 0.05$) words, and word-final V gestures were longer in stress-final words as compared to either stress-initial (S1, with $\beta = 32.1$, $SE = 1.9$, $p < 0.05$) or stress-medial (S2, with $\beta = 15.3$, $SE = 1.8$, $p < 0.05$) words.

The analysis of pairwise comparisons pointed to distinctions between the two unstressed conditions, as was the case for C gestures as well. V gestures of syllables adjacent to stressed ones were longer than those of syllables further away from the stress, indicating that stress has both anticipatory and spillover effects (see also Katsika & Tsai, 2019). To be more precise, word-initial V gestures were longer in stress-medial words as opposed to stress-final words ($S2 > S3$ for V1, with $\beta = 22.3$, $SE = 1.8$, $p < 0.05$), and word-final V gestures were longer in stress-medial words than in stress-initial words ($S2 > S1$ for V3, with $\beta = 16.8$, $SE = 1.7$, $p < 0.05$). Word-medial V gestures, which, when unstressed, are by default adjacent to a stressed syllable, were slightly longer in stress-final words as compared to stress-initial ones ($S3 > S1$ for V2, with $\beta = 5$, $SE = 1.7$, $p < 0.05$). This latter pattern indicates that anticipatory effects are stronger than spillover ones on V gestures. As a reminder, C gestures showed the opposite pattern, i.e., stronger spillover than anticipatory effects. These two patterns in combination suggest that closest proximity matters: onset V gestures of medial syllables, although in-phase (synchronous) with their syllable's onset C gesture (cf. Katsika, 2014), are immediately adjacent to final stress by virtue of their longer duration. The C gesture, on the other hand ends much earlier than the V gesture, and hence does not directly neighbor final stressed syllables. The distinctions between stress positions

held regardless of focus types for all V gestures, as the analysis of the first interaction listed in Table 4 showed ($p < 0.05$ for all pairwise comparisons).

Applying a syntagmatic, and not a paradigmatic, approach on these pairwise comparisons, we found that regardless of stress position, word-medial and word-final V gestures were not significantly different from each other in terms of duration, despite their different vowel quality (/i/ for V2 and /e/ for V3), except in stress-final words in which final V gestures were longer than the medial ones by a marginally significant difference of 25 ms ($SE = 12$, $p = 0.09$). This equation in duration between medial and final V gestures could be attributed to stress proximity effects. Moreover, word-initial V gestures were the longest in stress-initial words (S1: $V1 > V2$, with $\beta = 23.2$, $SE = 7.4$, $p < 0.05$ and $V1 > V3$, with $\beta = 25.4$, $SE = 9.3$, $p < 0.05$), and the shortest in stress-medial and stress-final words (S2: $V2 > V1$, with $\beta = 38.3$, $SE = 7.4$, $p < 0.05$ and $V3 > V1$, with $\beta = 22.6$, $SE = 9.2$, $p < 0.05$; S3: $V2 > V1$, with $\beta = 35.3$, $SE = 7.4$, $p < 0.05$ and $V3 > V1$, with $\beta = 60.2$, $SE = 9.3$, $p < 0.05$), suggesting that spillover effects are stronger than anticipatory ones.

Thus, the durational patterns of V gestures support an account of prominence in which stress effects are not local to the stressed syllable but extend to both preceding and following parts of the word, with spillover effects being larger than anticipatory ones. As a result, initial V gestures, when unstressed, are the shortest V gestures in the word, which could possibly function as an effective cue for word demarcation.

3.1.2.3. Effects of focus and accent.: A three-way interaction effect was detected for the factors of focus type, stress position and syllable affiliation (Table 4, interaction 3). The associated pairwise comparisons revealed that focus type had durational effects, although not connected to the stressed syllable. The duration of the word's final V gesture encoded focus type regardless of which syllable of the word was stressed (see Fig. 10). Word-final V gestures were the longest in words under broad focus. This pattern held for all stress positions (BF > NF for V3 in S1: $\beta = 24.3$, $SE = 2.8$, $p < 0.05$, S2: $\beta = 27.7$, $SE = 2.37$, $p < 0.05$, and S3: $\beta = 26.2$, $SE = 3$, $p < 0.05$; BF > UF for V3 in S1: $\beta = 32.7$, $SE = 2.7$, $p < 0.05$, S2: $\beta = 30$, $SE = 2.7$, $p < 0.05$, and S3: $\beta = 33.1$, $SE = 2.7$, $p < 0.05$). In stress-initial and stress-final words, an additional distinction was observed between narrow focus and unfocused conditions, with V gestures being longer in the former (NF > UF for V3 in S1: $\beta = 8.5$, $SE = 2.3$, $p < 0.05$; NF > UF in S3: $\beta = 6.9$, $SE = 2.4$, $p < 0.05$). Stress-medial words did not present this distinction.

Word-medial V gestures did not distinguish focus types, but they did differentiate between accented/focused and de-accented/unfocused conditions, being the shortest when unfocused (UF < BF for V2 in S1: $\beta = 13.3$, $SE = 2.7$, $p < 0.05$, S2: $\beta = 16.1$, $SE = 2.6$, $p < 0.05$, S3: $\beta = 6.7$, $SE = 2.8$, $p < 0.05$; UF < NF for V2 in S1: $\beta = 14.6$, $SE = 2.3$, $p < 0.05$, S2: $\beta = 14.4$, $SE = 2.3$, $p < 0.05$, S3: *n.s.*). Finally, word-initial V gestures did not encode accent or focus. The only exception was stress-medial words, in which the narrow focus condition was longer than the unfocused one (NF > UF for V1 in S2: $\beta = 9.4$, $SE = 2.4$, $p < 0.05$).

The analysis of the interaction effects between focus type, syllable position and stiffness (Table 4, interaction 2) on one hand, and between focus type with boundary type and stress

position (Table 4, interaction 3) on the other hand, further confirmed that the three focus types were distinguished from each other only in word-final V gestures (V3: BF > NF, with $\beta = 14.8$, $SE = 1.38$, $p < 0.05$; BF > UF, with $\beta = 17.7$, $SE = 1.6$, $p < 0.05$; NF > UF only in word-medial positions (W), with $\beta = 8.9$, $SE = 1.7$, $p < 0.05$). In word-initial and word-medial V gestures, each of the two focused conditions was longer than the unfocused one ($p < 0.05$, except for the comparison between NF and UF for phrase-final V1 gestures, which was not significant), but the two focused conditions were not differentiated from each other.

To conclude, it is the position of the V gesture within the word, and not whether its syllable is stressed or not, that determines whether gestural duration encodes focus and accent. The most informative position in this respect is the word-final one, which differentiates among the three types of focus. Note that these patterns go against Hypothesis II in several ways. First, it is not the stressed syllables that undergo the effects, and consequently, stressed syllables do not reflect a hierarchy of prominence. Second, contrary to Hypothesis II, which stipulated that it would be narrow focus that would present the longest durations, the emerging patterns showed that broad focus was the longest, narrow focus the next longest and unfocused the shortest.

3.1.2.4. Effects of phrase boundaries.: Stress position was involved in several interactions with boundary type. Importantly, boundary-related lengthening extended further away from the boundary in stress-initial and stress-medial words as opposed to stress-final ones, confirming previous findings, reported in Katsika (2016) (Table 4, interaction 4; see Fig. 11).

In stress-initial and stress-medial words, both medial and final V gestures lengthened due to an upcoming phrase boundary (V2 in S1: IP > W, with $\beta = 11$, $SE = 2.4$, $p < 0.05$; V2 in S2: IP > W, with $\beta = 10.5$, $SE = 2.4$, $p < 0.05$; V3 in S2: IP > W, with $\beta = 9$, $SE = 2.5$, $p < 0.05$; V3 in S1: IP > W, with $\beta = 35.6$, $SE = 2.5$, $p < 0.05$; V3 in S2: IP > W, with $\beta = 25.4$, $SE = 2.4$, $p < 0.05$). In stress-final words, it was only the final V gesture that presented boundary-related lengthening (V3 in S3: IP > W, with $\beta = 31.6$, $SE = 2.6$, $p < 0.05$). In all stress positions, initial V gestures underwent boundary-related shortening (V1 in S1: W > IP, with $\beta = 6$, $SE = 2.5$, $p < 0.05$; V1 in S2: W > IP, with $\beta = 5.8$, $SE = 2.3$, $p < 0.05$; V1 in S3: W > IP, with $\beta = 9$, $SE = 2.5$, $p < 0.05$).

Overall, at the mean value of stiffness ($\text{mean}_{\text{stiffness}} = 16.9$ - Hz), V gestures of phrase-final words were the longest when stress was medial, the next longest when stress was initial and the shortest when stress was final (IP: S2 > S1, with $\beta = 3.5$, $SE = 1$, $p < 0.05$; S2 > S3, with $\beta = 12.5$, $SE = 1$, $p < 0.05$; S1 > S3, with $\beta = 9$, $SE = 1.1$, $p < 0.05$) (Table 4, interaction 5). Phrase-medially, the distinction between medial stress and either initial or final stress remained (IP: S2 > S1, with $\beta = 7.1$, $SE = 1.7$, $p < 0.05$; S2 > S3, with $\beta = 9$, $SE = 1.7$, $p < 0.05$), but the distinction between initial and final stress disappeared.

3.1.3. Summary of results on gestural duration—Table 5 summarizes the results on gestural duration.

3.2. Displacement

3.2.1. Displacement of C gestures—The minimal adequate model of displacement for C gestures included random intercepts and random slopes for syllable affiliation per speaker, and the significant effects of the fixed factors summarized in Table 6.

3.2.1.1. Relationship between displacement and peak velocity: Displacement of C gestures was found to increase linearly with peak velocity with a slope of 0.34 mm per cm/s, as shown in Fig. 12(a) (Fig. 12(b) is discussed in Section 3.2.2.1).

3.2.1.2. Effects of stress: Stress affected the amount of displacement of C gestures, as illustrated in Fig. 13.

At mean peak velocity value ($\text{mean}_{\text{peak velocity}} = 11 \text{ cm/sec}$), initial C gestures were the largest when stressed (in C1: $S1 > S2$, with $\beta = 0.4$, $SE = 0.1$, $p < 0.05$, and $S1 > S3$, with $\beta = 0.6$, $SE = 0.1$, $p < 0.05$), and did not undergo any displacement effect when stress was on a subsequent syllable (i.e., the medial or final syllable) (Table 6, interaction 1). When checking each focus type separately (Table 6, interaction 3), only the distinction between initial and final stress remained ($S1 > S3$ in BF ($p < 0.05$), NF ($p < 0.05$) and UF (*m.s.*); see Fig. 14). On the other hand, both medial and final C gestures were the largest when stress was on the preceding syllable (in C2: $S1 > S2$, with $\beta = 5.9$, $SE = 0.1$, $p < 0.05$, and $S1 > S3$, with $\beta = 8.7$, $SE = 0.1$, $p < 0.05$; in C3: $S2 > S3$, with $\beta = 0.6$, $SE = 0.1$, $p < 0.05$, and $S2 > S1$, with $\beta = 1.2$, $SE = 0.1$, $p < 0.05$), and the next largest when they belonged to the stressed syllable (in C2, $S2 > S3$, with $\beta = 2.8$, $SE = 0.1$, $p < 0.05$; in C3, $S3 > S1$, with $\beta = 0.7$, $SE = 0.1$, $p < 0.05$). The patterns observed in the medial C gesture held regardless of focus type (Table 6, interaction 3; $p < 0.05$ for all pairwise comparisons; see Fig. 14). Final C gestures, however, presented different patterns depending on focus type. Medial stress was related to the largest displacement solely in broad focus (C3: $S2 > S1$, $S3$ in BF, with $p < 0.05$ for all comparisons), while initial stress was related to the smallest displacement solely in narrow focus (C3: $S2$, $S3 > S1$ in NF, with $p < 0.05$ for all comparisons). When unfocused, final C gestures distinguished only between initial and medial stress positions ($S2 > S1$ in UF, with $p < 0.05$).

Taking the syntagmatic approach to stress effects on displacement revealed that initial and medial stress positions affected displacement, while final stress position did not (Table 6, interaction 1). In stress-initial and stress-medial words, the largest displacement was observed in medial C gestures ($C2 > C1$ in S1, with $\beta = 9.7$, $SE = 0.9$, $p < 0.05$, and in S2, with $\beta = 4.2$, $SE = 0.9$, $p < 0.05$; $C2 > C3$ in S1, with $\beta = 10.9$, $SE = 0.9$, $p < 0.05$, and in S2, with $\beta = 3.8$, $SE = 0.9$, $p < 0.05$). In stress-initial words, initial C gestures were also slightly larger than the final ones ($C1 > C3$ in S1, with $\beta = 1.4$, $SE = 0.2$, $p < 0.05$). These patterns held in all focus types (i.e., in S1, $C2 > C1$, C3 and $C1 > C3$, with $p < 0.05$ for all pairwise comparisons; in S2, $C2 > C1$, C3, with $p < 0.05$ for all pairwise comparisons except for $C2 > C3$ in UF, which was marginal significant; finally, there was no significant comparisons in S3) (Table 6, interaction 3; see Fig. 14).

To conclude, the paradigmatic and syntagmatic analyses of C gesture displacement converge in that stress, in addition to affecting the stressed syllable, as predicted by Hypothesis IIa, has spillover effects. In fact, the peak of the effect is on syllable following the stressed one.

3.2.1.3. Effects of focus and accent.: The analysis of the interaction of focus type with stress position and syllable affiliation (Table 6, interaction 3; shown in Fig. 14) did not detect a systematic effect of focus and/or accent on the displacement of C gestures. The hypothesis that C gestures of stressed syllables would become larger when accented did not hold (Hypotheses IIb). It was only stress-medial C gestures that presented larger displacement when accented (C2 gestures of S2 words: BF > UF, with $\beta = 1.1$, $SE = 0.2$, $p < 0.05$, and NF > UF, with $\beta = 1.5$, $SE = 0.2$, with $p < 0.05$ for all comparisons). In broad focus, it was also the C gesture of the post-stress syllable that underwent spatial amplification (C3 gesture of S2 words: BF > UF, with $\beta = 0.8$, $SE = 0.2$, $p < 0.05$), suggesting the presence of spillover effects. This relation between penultimate stress and focus marking held regardless of prosodic boundary type (In S2 words, NF > UF in IP and W; BF > UF in IP (there was no BF condition in W), with $p < 0.05$ for all pairwise comparisons; Table 6, interaction 6).

In addition, analysis of interaction 2 in Table 6 showed that at mean peak velocity ($mean_{pv} = 11$ cm/sec), medial and final C gestures distinguished between focused/accented and unfocused/unaccented conditions, with words under broad and narrow focus involving larger medial and final gestures than their unfocused counterparts by an average of 0.5 mm ($SE = 0.1$). All pairwise comparisons were significant ($p < 0.05$), except for the comparison between NF and UF for C3, which was marginally significant. In all focus conditions, medial C gestures were the largest (C2 > C1, C3 in BF, NF and UF; $p < 0.05$ for all pairwise comparisons). Lastly, C gestures were overall the largest in stress-initial words, followed by stress-medial words regardless of focus type and boundary type (S1 > S2 > S3 under BF, NF or UF focus and in both IP and W positions, with $p < 0.05$ for all pairwise comparisons; Table 6, interaction 6).

We could thus conclude that, Hypothesis II (b and c) does not hold, since, similarly to durational effects, neither focus nor accent further amplifies the displacement of stressed C gestures. Instead, phrasal prominence exerts effects of small magnitude over a greater domain, possibly the word as a whole, with stress position being relevant as to how these effects develop across their scope.

3.2.1.4. Effects of phrase boundaries.: As Fig. 15 shows, boundary type interacted with stress position in determining the displacement profile of C gestures (Table 6, interaction 5).

Stress-initial words presented larger medial C gestures phrase-medially as opposed to phrase-finally (in S1, C2 was larger in W > IP, with $\beta = 1.4$, $SE = 0.2$, $p < 0.05$), stress-medial words had larger final C gestures phrase-finally as opposed to phrase-medially (in S2, C3 was larger in IP > W, with $\beta = 0.7$, $SE = 0.2$, $p < 0.05$), and stress-final words did not distinguish between the two prosodic positions.

Analysis of interaction 4 in Table 6 clarified that, regardless of stress position, at average peak velocity value ($mean_{pv} = 11$ cm/sec), medial C gestures were larger in phrase-medial positions (C2: W > IP, with $\beta = 0.8$, $SE = 0.1$, $p < 0.05$) and final C gestures in phrase-final ones (C3: IP > W, with $\beta = 0.3$, $SE = 0.1$, $p < 0.05$). Initial C gestures did not encode information pertaining to boundary type. When focus type was considered in combination with stress position (Table 6, interaction 6), it was further found that unfocused stress-initial words had in general larger C gestures in phrase-medial positions as opposed to phrase-final ones (S1 in UF: W > IP, with $\beta = 0.7$, $SE = 0.1$, $p < 0.05$).

These patterns confirm our previous conclusion that prominence-induced displacement effects peak on the syllable that immediately follows the stressed one. The effects surface as such phrase-medially, but phrase-finally, they also interact with boundary-related effects that enlarge the boundary-adjacent gestures regardless of stress position (see Hypothesis IV).

3.2.2. Displacement of V gestures—The minimal adequate model of displacement for V gestures included random intercepts and random slopes for syllable affiliation per speaker. Table 7 lists the significant interactions of the fixed factors.

3.2.2.1. Relationship between displacement and peak velocity.: As expected, displacement of V gestures increased linearly with peak velocity with a slope of 0.5 mm per cm/sec (see Fig. 12b).

3.2.2.2. Effects of stress.: At their average peak velocity ($mean_{pv} = 17$ mm/ms), V gestures were the largest when stressed and the next largest when adjacent to stress (Table 7, interaction 1). As Fig. 16 illustrates, initial V gestures were the largest in stress-initial words and the smallest in stress-final words (V1: S1 > S2, S3 and S2 > S3, with $p < 0.05$ for all pairwise comparisons). In parallel, final V gestures were the largest in stress-final words and the smallest in stress-initial words (V3: S3 > S1, S2 and S2 > S1, with $p < 0.05$ for all pairwise comparisons). Medial V gestures were the smallest in stress-final words (V2: S1, S2 > S3, with $p < 0.05$ for all pairwise comparisons), while the slightly larger displacement they presented in stress-medial words as opposed to stress-initial words was marginally significant (V2: S2 > S1, $p = 0.06$). These patterns held for all focus types (Table 7, interaction 4; $p < 0.05$ for all comparisons except for the V3 comparison between S1 and S2 in BF, which was not significant).

Taken together, these findings further support the conclusion that kinematic effects of stress extend beyond the stressed syllable, with the spillover direction being stronger than the anticipatory one.

The paradigmatic approach reinforced the conclusion that gestures become larger when stressed (Table 7, interaction 1). Initial, medial and final V gestures were the largest in stress-initial, stress-medial and stress-final words respectively (In S1: V1 > V2, V3; in S2: V2 > V1; in S3: V3 > V1, V2, with $p < 0.05$ for all pairwise comparisons except between V2 and V3 in S2 which was not significant). Examining these patterns by focus type (Table 7, interaction 4) revealed no systematic interaction with focus, as shown in Fig. 17. Stress-final words presented larger displacement in stressed V gestures regardless of

focus type (S3: $V3 > V1$, $V2$ in BF, NF, and UF, with $p < 0.05$ for all pairwise comparisons). When under broad focus, stress-final words also presented a finer distinction between their two unstressed V gestures, with the V gesture adjacent to stress being larger than the more remote one (S3 in BF: $V2 > V1$, $p < 0.05$). In stress-initial and stress-medial words, on the other hand, effects of stress on stressed V gestures were dependent on focus type. In stress-initial words, stressed V gestures were larger than the unstressed ones when the word was unfocused (S1 in UF: $V1 > V2$, $V3$, with $p < 0.05$). In narrow focus, only the distinction between initial and final V gestures held (S1 in NF: $V1 > V3$, $p < 0.05$), while no comparison was significant in broad focus. Finally, in stress-medial words their medial V gestures were larger than their initial ones only when they were focused ($V2 > V1$ in BF and NF, with $p < 0.05$). These patterns can be accounted for if we assume that presence of accent/focus affects a domain larger than the stressed syllable, often eliminating the distinctions between stressed and unstressed V gestures within a word.

In sum, both the syntagmatic and the paradigmatic analyses suggest that V gestures are larger when stressed as opposed to unstressed, confirming Hypothesis IIa. Stress has strong spillover and weaker anticipatory effects. Similarly to duration, effects of accent and/or focus on displacement of V gestures are neither localized nor intensified on the stressed syllable, rejecting Hypotheses IIb and IIc. This latter conclusion received further support from the analyses reported in subsection 3.2.2.3.

3.2.2.3. Effects of focus and accent.: As Fig. 17 shows, stressed syllables did not encode focus type systematically (Table 7, interaction 1). Stressed V gestures of stress-final words were the only ones that distinguished among all three focus types tested here, showing the largest displacement in broad focus, the next largest displacement in narrow focus and the smallest displacement when unfocused ($V3$ in S3: BF $>$ NF $>$ UF, with $p < 0.05$ for all comparisons). Stressed V gestures of stress-medial words were larger when focused/ accented as opposed to unfocused/de-accented ($V2$ in S2: BF, NF $>$ UF, with $p < 0.05$ for all comparisons), while stressed V gestures of stress-initial gestures did not present any effect of focus/accnt.

On the basis of these results, the hypothesis that stressed V gestures become even larger when accented and/or focused is partially rejected (Hypotheses IIb and c). Instead, the hypothesis needs to be altered in order to reflect the finding that the position of the stressed V gesture within the word matters in how focus type is encoded. This conclusion is further confirmed by the interaction effect between focus type and syllable affiliation (Table 7, interaction 2), according to which only final V gestures distinguished among all three types of focus, being the largest under broad focus, the next largest under narrow focus and the smallest when unfocused ($V3$: BF $>$ NF $>$ UF, with $p < 0.05$ for all comparisons). Initial V gestures were larger under narrow focus and medial V gestures were larger under broad focus as opposed to being unfocused, but did not encode any other focus type information ($V1$: NF $>$ UF, $p < 0.05$; $V2$: BF $>$ UF, $p < 0.05$).

Finally, although words distinguished between focused and unfocused V gestures, regardless of stress position (BF, NF $>$ UF in S1, S2 and S3, with $p < 0.05$ for all comparisons), it was only stress-initial words that further distinguished between the two focused conditions

(BF > NF, $p < 0.05$; Table 7, interaction 3). In parallel, in words under broad focus, V gestures were the largest when stress was initial, next largest when stress was medial and the smallest when stress was final. Narrow-focused and unfocused words, on the other hand, did not distinguish between initial and medial stress positions, the V gestures of which were equally large to each other and larger than those of stress-final words (BF: $S1 > S2 > S3$; NF and UF: $S1, S2 > S3$, with $p < 0.05$ for all comparisons; Table 7, interaction 3). With respect to phrasal position, V gestures were larger when focused/accented as opposed to unfocused/de-accented, regardless of boundary type (IP: BF, NF > UF; W: NF > UF, with $p < 0.05$ for all comparisons; Table 7, interaction 6).

The analyses reported here reinforce the conclusion that it is mainly position of gestures and stress within the word that matter in how accent and/or focus is kinematically marked, and not the stressed syllable itself as put forward in Hypothesis II.

3.2.2.4. Effects of phrase boundaries.: The type of prosodic boundary interacted separately with stress position (Table 7, interaction 5) and syllable affiliation (Table 7, interaction 7), as shown in Fig. 18. Specifically, initial V gestures were larger in phrase-medial words as opposed to phrase-final ones, while for final V gestures the opposite pattern was true (V1: W > IP, $p < 0.05$; V3: IP > W, $p < 0.05$; Table 7, interaction 7). At mean peak velocity ($mean_{pv} = 17$ cm/sec), V gestures of phrase-final stress-initial words were larger than their phrase-medial counterparts (S1: IP > W, $p < 0.05$; Table 7, interaction 5). In general, the displacement of V gestures of phrase-final words were the largest when stress was initial, the next largest when stress was medial and the smallest when stress was final (IP: $S1 > S2 > S3$, with $p < 0.05$ for all comparisons; Table 7, interaction 5).

In conclusion, as expected, gestures became larger when adjacent to IP boundary. This amplification was preceded by attenuation of initial V gestures. As predicted by Hypothesis III, position of stress further interacted with these effects: the further the stress was from the boundary the larger was the displacement.

In general, from the analyses reported so far, initial and medial syllables emerge as loci of maximum prominence-related temporal and spatial effects. This is possibly because stress marking involves significant spillover effects, which in stress-initial and stress-medial words have the potential of being expressed within word boundaries. It also suggests that prominence- and boundary-related effects do not function additively.

3.2.3. Summary of results on gestural displacement—Table 8 summarizes the results on gestural displacement.

3.3. Peak velocity

3.3.1. Peak velocity of C gestures—The minimal adequate model of peak velocity for C gestures included random intercepts and random slopes for syllable affiliation per speaker. Table 9 lists the significant effects of the fixed factors.

3.3.1.1. Relationship between peak velocity and displacement.: Peak velocity of C gestures increased linearly with displacement with a slope of 1.8 cm/sec per mm, shown in Fig. 19(a) (Fig. 19(b) is discussed in Section 3.3.2.1).

3.3.1.2. Effects of stress.: As Fig. 20 demonstrates, there was no systematic effect of stress on the peak velocity of stressed C gestures (Table 9, interaction 1).

At average C gesture displacement ($mean_{displacement} = 5.6$ - mm), medial C gestures presented the predicted pattern of being faster when stressed (C2: S2 > S1, S3, $p < 0.05$), but initial and final C gestures did not, contradicting Hypotheses I and IIa. Word-initial gestures did not show any effect of stress, while final gestures were the fastest when stress was initial, the next fastest when stress was final and the slowest when stress was medial (C3: S1 > S3 > S2, with $p < 0.05$ for all comparisons). These patterns held regardless of focus type (Table 9, interaction 4: $p < 0.05$ for all comparisons, except for S3 > S2 of C3 in BF, which was marginally significant, with $p = 0.06$).

The syntagmatic approach to stress relations further rejected Hypotheses I and IIa, confirming that stress did not have systematic effects on stressed C gestures. In stress-initial words, it was the final C gesture that was the fastest (S1: C3 > C1, C2, $p < 0.05$), while initial C gestures were faster than medial (S1: C1 > C2, $p = 0.05$). Stress-medial words did not present any significant comparison between their gestures, and stress-final words had equally fast initial and final gestures, which were in turn faster than their medial gestures (S3: C1, C3 > C2, with $p < 0.05$ for all comparisons). When looking at each focus type separately (Table 9, interaction 4), an additional distinction was detected for unfocused stress-final words. In particular, when unfocused, stressed C gestures of stress-final words were faster than the unstressed ones (S3 in UF: C3 > C2, C1, with $p < 0.05$ for all comparisons), as shown in Fig. 21.

On the basis of our analyses, we can thus conclude that when C gesture peak velocity is considered independently from displacement, it is not consistently affected by stress. Instead, the velocity profile of the word is different for each stress position. These findings contradict the velocity-related statements of Hypotheses I and II.

3.3.1.3. Effects of focus and accent.: Neither did the analysis of the three-way interaction between the factors of focus type, stress position and syllable affiliation (Table 9, interaction 4) show any systematic effect of focus type on the peak velocity of C gestures, stressed or not. This is also apparent in Fig. 21.

Analysis of the interactions of focus type with either stress position (Table 9, interaction 3) or syllable affiliation (Table 9, interaction 2) revealed that any effects of focus on peak velocity of C gestures could not be generalized across stress profiles and syllables, invalidating Hypothesis III. At average C gesture displacement ($mean_{disp} = 5.6$ mm), it was only stress-final words that encoded focus type, with C gestures of words under narrow focus being faster than their unfocused counterparts (S3: NF > UF, $p < 0.05$). In parallel, when words were unfocused their C gestures were the fastest in medial stress, the next

fastest in initial stress, and the slowest in final stress (UF: $S2 > S1 > S3$, with $p < 0.05$ for all comparisons).

Similarly, the position of the gesture in the word affected the number of types of focus encoded. Medial C gestures differentiated between focused and unfocused conditions, with both broad- and narrow-focused words having faster medial C gestures than their unfocused counterparts (C2: BF, NF $>$ UF, with $p < 0.05$ for all comparisons). However, initial and final C gestures distinguished the unfocused condition from narrow focus and broad focus respectively (C1: NF $>$ UF; C3: BF $>$ UF, with $p < 0.05$ for all comparisons). Finally, in unfocused words, final C gestures were faster than initial and medial C gestures (UF: C3 $>$ C1, C2, with $p < 0.05$ for all comparisons). According to the analysis of the interaction between focus type and boundary (Table 9, interaction 7), C gestures under narrow focus differed from their unfocused counterparts only in phrase-final positions (IP: NF $>$ UF, $p < 0.05$).

Like stress, focus type did not affect peak velocity systematically, contradicting the velocity-related statements of Hypothesis I and II. Instead, different positions of stress and syllabic affiliations of gestures patterned differently. Notably, despite these differences, significant comparisons presented the expected direction of the effects, with focused conditions being faster than unfocused ones.

3.3.1.4. Effects of phrase boundaries.: One of the main expected differences between boundary and prominence marking is that gestures slow down at boundaries, but become faster under prominence. Our analysis of the three-way interaction among boundary, stress position and syllable affiliation (Table 9, interaction 6) confirmed that gestures closer to the IP boundary were slower as compared to their phrase-medial counterparts. However, this pattern did not hold across the board, but was present in the medial C gesture of stress-final words and the final C gesture of stress-medial words (C2 in S3: W $>$ IP; C3 in S2: W $>$ IP, with $p < 0.05$ for all comparisons). C gestures of stress-initial words did not present any boundary-related effect on the dimensions of peak velocity. These effects are illustrated in Fig. 22.

The interaction between boundary type and stress position (Table 9, interaction 5) was such that only stress-medial words presented a general pattern of slower C gestures phrase-finally as opposed to phrase-medially (S2: W $>$ IP). The effect was marginally significant with $p = 0.06$. This analysis also revealed that C gestures were overall slower in stress-final words as opposed to either stress-initial or stress-medial words in both phrase-medial and phrase-final positions (S1, S2 $>$ S3 in both IP and W, with $p < 0.05$ for all comparisons except for S1 $>$ S3 in IP, for which $p = 0.06$ (*m.s.*)).

3.3.2. Peak velocity of V gestures—The minimal adequate model of peak velocity for V gestures included random intercepts and random slopes for syllable affiliation per speaker. Table 10 lists the significant effects of the fixed factors.

3.3.2.1. Relationship between peak velocity and displacement.: As Fig. 19 b) illustrates, peak velocity of V gestures increased linearly with displacement with a slope of 1.1 cm/sec per mm.

3.3.2.2. Effects of stress.: There were no systematic effects of stress on the peak velocity of V gestures, as shown in Fig. 23 (Table 10, interaction 1). At mean V gesture displacement ($mean_{displacement} = 10.5$ mm), initial V gestures were the fastest when stress was final, and the next fastest when stress was medial (V1: S3 > S2 > S1, with $p < 0.05$ for all comparisons). Final V gestures, on the other hand, were the fastest in stress-initial words (V3: S1 > S2, S3, with $p < 0.05$ for both comparisons), but other stress positions did not present any distinctions. Finally, peak velocity of medial V gestures was not affected by stress.

Examining the interaction of these effects with focus type (Table 10, interaction 4; Fig. 24) showed that the patterns observed in initial V gestures remained intact by focus type (V1: S3 > S2 > S1 in BF, NF and UF, with $p < 0.05$ for all pairwise comparisons). Medial V gestures were the slowest when stressed as opposed to unstressed regardless of focus type (V2: S1, S3 > S2 in BF, NF and UF, with $p < 0.05$ for all pairwise comparisons). Under narrow focus, medial V gestures were also faster in initial as opposed to final stress (V2: S1 > S3 in NF, with $p < 0.05$). Final V gestures were faster in stress-initial words than in stress-medial words in all three types of focus (V3: S1 > S2, in BF, NF and UF, with $p < 0.05$ for all pairwise comparisons). When narrow-focused or unfocused, final V gestures were also faster in stress-initial words than in stress-final words (V3 in NF and UF: S1 > S3, with $p < 0.05$ for all pairwise comparisons). Finally, under narrow focus, final V gestures of stress-medial words were faster than their stress-final counterparts (V3 in NF: S2 > S3, $p < 0.05$).

On the paradigmatic dimension, only stress-final words distinguished among their V gestures (Table 10, interaction 1): initial V gestures were the fastest and final V gestures the slowest (S3: V1 > V2 > V3, with $p < 0.05$ for both comparisons). In the other stress positions, V gestures were roughly of same peak velocity, except stress-initial words' medial V gestures, which were faster than their initial V gestures (S1: V2 > V1, $p < 0.05$). These patterns remained stable across focus types ($p < 0.05$ for all comparisons; Table 10, interaction 4).

3.3.2.3. Effects of focus and accent.: Contrary to Hypothesis II (b and c), stressed V gestures of focused words were not faster than their unfocused counterparts (Table 8, interaction 4), as demonstrated in Fig. 24. In fact, focus type was not related to any systematic effects of peak velocity on any V gesture (Table 8, interactions 2 and 3). The only comparisons that were significant were the following: V gestures of stress-final words were faster when unfocused as opposed to narrow-focused (S3: UF > NF, $p < 0.05$). In general, in the unfocused and broad focus conditions, V gestures were the fastest in stress-final words, whereas under narrow focus, V gestures were the slowest in stress-medial words (BF and UF: S3 > S1, S2; NF: S1, S3 > S2, with $p < 0.05$ for all comparisons).

3.3.2.4. Effects of phrase boundaries.: Similarly to focus and/or accent, prosodic boundaries did not have the effects expected by Hypothesis IV on peak velocity. As a

reminder, we expected V gestures to slow down by phrase boundaries. Our analyses did not find any systematic effects of stress on peak velocity of phrase-final, or other, gestures (Table 10, interactions 5, 6 and 7). Fig. 25 demonstrates the peak velocity profile of V gestures by boundary type. There were however interaction effects between type of boundary and stress position. V gestures were the slowest in stress-medial words, regardless of boundary type, while, only in phrase-final positions, V gestures were also the fastest when occurring in stress-final words (IP: $S3 > S1 > S2$; W: $S1, S3 > S2$, with $p < 0.05$ for all comparisons; Table 8, interaction 6). Medial V gestures were overall faster than initial V gestures in phrase-medial positions (W: $V2 > V1$, $p < 0.05$; Table 10, interaction 5).

3.3.3. Summary of results on gestural peak velocity—Table 11 summarizes the results on gestural peak velocity.

4. Discussion

4.1. General kinematic profile of prominence

To summarize, a general pattern arises in our data: prominent gestures in Greek are longer, larger and faster than their non-prominent counterparts, confirming our Hypothesis I, which was based on a substantial body of previous work (e.g., Beckman et al., 1992; Beckman & Edwards, 1994; Cho, 2005, 2006; de Jong, 1991, 1995; de Jong et al., 1993; Fowler, 1995; Harrington et al., 2000; Harrington et al., 1995; Mücke & Grice, 2014). Our analyses further clarify that a significant part of the spatio-temporal expansion that gestures undergo under prominence occurs independently of the relationship of duration with stiffness and of displacement with peak velocity. However, prominent gestures become faster by virtue of being larger, i.e., due to the linear relationship of peak velocity with displacement (see e.g., Munhall et al., 1985; Ostry & Munhall, 1985).

4.2. The hierarchy of prominence

Contrary to Hypothesis II, a hierarchy of prominence is not reflected in the kinematic information of syllables assigned stress. The spatiotemporal effects detected on these syllables encode lexical stress, but not pitch accent and/or focus. As a reminder, Hypothesis II expected both lexical stress and accentuation to affect the same kinematic dimensions of stressed gestures, but to differ in the degree of modulation, such that the higher the prominent unit in the hierarchy of prominence (i.e., accented > stressed > unstressed), the higher the degree of the effect (*stressed vs. unstressed*: e.g., Arvaniti, 1991, 2000; Baltazani, 2007; Botinis, 1982, 1989; Dauer, 1980; Fourakis et al., 1999; Nicolaidis, 2003; Nicolaidis & Rispoli, 2005, but not in Arvaniti, 1991, 2000; Dauer, 1980; see also Beckman & Edwards, 1994, Crystal & House, 1988, de Jong, 2004; Kelso et al., 1985; *accented vs. unaccented*: e.g., Baltazani & Jun, 1999; Botinis et al., 2002; Botinis & Bannert, 2003; Kastrinaki, 2003; but not in Botinis, Fourakis, & Katsaiti, 1995; Botinis et al., 2001a; Botinis et al., 2001b; see also: Beckman et al., 1992; Beckman & Edwards, 1994; Cambier-Langeveld & Turk, 1999; Cho, 2002, 2005, 2006; Cho & Keating, 2009; Cho & McQueen, 2005; de Jong, 1991, 1995, 2004; de Jong et al., 1993; Fowler, 1995; Harrington et al., 1995, 2000; Turk & Sawusch, 1997; Turk & White, 1999, but see Mücke & Grice, 2014). In our data, stressed gestures involve indeed more extreme movements than unstressed gestures,

but, contrary to predictions, presence of accent does not further expand the movements of stressed gestures.

This divergence from previously reported patterns might be related to the different conditions of prominence used by the different studies, rendering direct comparisons difficult. For instance, the current study obtains de-accented test words by placing the nuclear pitch accent earlier in the utterance by two words. On the other hand, studies like Cho (2004, 2005) and Beckman and Edwards (1994) contrasted accented test words to their de-accented counterparts by placing contrastive focus either on the test word or on one of its two immediate neighbor words. It is thus likely that a combination of type of focus on the nuclear accent with distance from it plays a role to the degree of de-accentuation a word undergoes, making for example the test words tested here less reduced than those in other studies (see discussion in Cho, Kim, & Kim, 2013). In the same vein, other studies used different prominence conditions. For example, de Jong compared pitch accented stressed syllables to unstressed syllables that consequently do not bear accent. However, the following two observations about the current dataset need to be kept in mind: (1) as shown in Fig. 1, in the de-accented conditions, F0 is extremely high during the nuclear pitch accent, while it compresses fully during the following test word, and (2) the part of the utterance separating the nuclear pitch accent from the test word is not necessarily long, specifically it counts five to six syllables.

Another dimension that prevents direct comparisons with previous research in general and in Greek specifically is that test words across studies vary in the position of stress. Given the contrastive and phonologically unpredictable nature of lexical stress in Greek, we did not expect its articulatory correlates to vary with its position within the word (Hypothesis III). Indeed, regardless of stress position, stressed gestures were overall longer, larger and faster than their unstressed counterparts. However, the minor effects that accent exerted, despite extending over the whole word, presented different profiles depending on stress position, possibly explaining why previous work has reached contradicting conclusions with respect to accentual lengthening in Greek (e.g., Baltazani & Jun, 1999; Botinis & Bannert, 2003; Botinis et al., 1995, 2001a, 2001b, 2002; Kastrinaki, 2003). For instance, the acoustic analysis of Greek in Botinis (1989) found robust evidence for stress-related lengthening, but inconsistent evidence for accentual lengthening.

Returning to the question of a hierarchy of prominence and the current set of data, despite not affecting the stressed syllable per se, accent causes some minimal gestural expansion across the whole word. These effects, albeit small in magnitude, correspond to the predicted hierarchy of prominence: they are stronger in accented words than in unaccented ones. The accented conditions examined here come from two focus types, meaning broad and narrow focus. With the exception of final V gesture duration, the effects of accent do not systematically reflect a consistent hierarchical order of these focus types, as was hypothesized based on previous findings on German and English (cf. German: Hermes et al., 2008; Mücke & Grice, 2014; Roessig & Mücke, 2019; English: Katsika et al., 2020). Depending on the dimension and gesture under question, it is either broad or narrow focus that shows the highest degree of kinematic modulation. For example, word-final V gestures are the longest in broad focus, but word-medial C gestures are the longest in narrow focus.

In what follows, we discuss some implications of the findings discussed in this section for the role of stress (Section 4.2.1) and the role of focus structure (Section 4.2.2) in prominence-marking.

4.2.1. The role of stress in prominence-marking—Taken together, the patterns found in this study suggest that in Greek, lexical stress affects the supra-laryngeal, kinematic profile of gestures. However, accent – at least, the types/degrees of accent studied here, i.e., broad or narrow focus placed some syllables away from the target – is mainly marked by pitch movements on the lexically stressed syllable, and does not regularly modulate further the kinematic development of the gestures co-occurring with those pitch movements. In this way, Greek might be similar to languages such as Arabic, in which their non-F0 phonetic correlates of stress do not become amplified when accented (cf. de Jong & Zawaydeh, 2002 for Arabic).

In this account, the patterns detected here lend support to the theoretical assumption that stressed syllables are the docking places for pitch accents. The longer, expanded gestures of stressed syllable allow for the additional laryngeal gesture or gestures of a pitch accent to be incorporated in the coordinative structure of the syllable, which is otherwise set at the mental lexicon level. Under this assumption, languages in which lengthening and articulatory expansion do not result from lexical stress or other lexically defined dimensions, such as phonological length, are more likely to induce these effects at the phrasal level in order to serve the incorporation of the pitch accent. Stress languages are thus expected to roughly present one of the following three hierarchical structures: (a) unstressed < stressed < accented, (b) unstressed < stressed/accented, and (c) unstressed/stressed < accented.

In the core of this proposal is the structural nature of stress; stress marks structural positions that bear the potential of spatio-temporal and/or tonal expansion, but the level or degree of prominence they express this potential is language-specific. Future research will need to assess this hypothesis by examining a wider range of degrees of prominence combined with controlled positioning of accents in the utterance, per the discussion developed in Section 4.2.

4.2.2. The role of focus structure in prominence-marking—Implicit in the hypothesis developed in Section 4.2.1 is the possibility that prominence is better represented as a continuum. It has traditionally been assumed that stress languages have three degrees of prominence, namely unstressed, stressed, and accented. We have already discussed how our data do not support the stressed vs. accented distinction for Greek and, based on this, reasons for which languages may differ in how they categorize stress kinematically (i.e., with unstressed, with accented, or between unstressed and accented). Recent work, however, has also provided evidence for finer distinctions within the accented category. This work has found phonetic correlates of focus structure in stressed syllables (e.g., broad vs. narrow vs. contrastive focus) instead of solely accentuation (unaccented vs. accented), pointing thus to more than three degrees of prominence (Hermes et al., 2008; Katsika et al., 2020; Mücke & Grice, 2014; Roessig & Mücke, 2019).

Expanding thus on the hypothesis proposed in Section 4.2.1, it is possible that the number of degrees of prominence that are phonetically encoded is greater than three, with the exact number being determined on a language-specific (if not speaker-specific) basis. Under this view, it might be the case that Greek nuclear pitch accents are not accompanied by kinematic modifications additional to the ones induced by stress, unless the accents denote types of focus and communicational functions (e.g., contrastive focus) that are connected to higher degrees of prominence than those denoted by broad and narrow focus (which are studied here). Although future research has yet to directly assess cross-linguistic comparisons in order to determine the hierarchical structure of prominence and its language-dependent dimensions, assuming cross-linguistic differences with respect to the number of phonetically marked degrees of prominence makes sense. This is because languages are known to differ in how they use phonetic correlates to mark stress and accent (see Fletcher, 2010 for an overview).

Section 4.2.1 argues that one of the roles of prominence-related articulatory expansion is to assist the incorporation of phrasal tonal events into the coordinative structure of the utterance. Extending this hypothesis, we could also assume that the height of F0 and/or the type of pitch accent (e.g., H*, L + H*, etc.) might be factors determining -or at least covarying with- the degree of kinematic modulation that accentuation induces. For instance, in Roessig and Mücke (2019), the increase in kinematic measures across focus types is accompanied by increase in F0 height. In parallel, there might be a correspondence between degree of prominence a pitch accent denotes and the type of pitch movements it involves. Intonational events are considered morphemes that encode pragmatic information, such as focus type (see review in Prieto, 2015). For example, in Greek declarative constructions, broad focus and narrow focus are marked by H* and L + H* pitch accents respectively (Arvaniti & Baltazani, 2005; Lohfink, Katsika, & Arvaniti, 2019). However, intonational meaning is currently understudied, and the connection between pitch accent types and their meaning function is not yet solidified (see e.g., Prieto, 2015). Future work should consider the role of intonational meaning along with focus structure in determining the hierarchy of prominence and its phonetic manifestation on both F0 and kinematic parameters.

4.3. Scope of stress-related effects

The current study controlled for position of stress in the word in order to examine the kinematic profile of stress across all possible stress positions in Greek. This methodological approach showed that stressed gestures are consistently longer and larger (and thus faster as well) than their unstressed counterparts, regardless of stress position. This is true not only for the V gestures of stressed syllables, but for their C gestures as well (only onset Cs were tested here), as indicated by studies in Greek (Arvaniti, 2000; Botinis et al., 2001a, 2001b), and other languages as well (e.g., Bombien et al., 2010, Cho & Keating, 2009). Importantly, this method also revealed finer distinctions between unstressed gestures. Gestures adjacent to stress were strengthened more as compared to gestures remote from stress.

On the basis of these distinctions, we could argue that stress is not binary, simply separating between stressed and unstressed domains. Instead, there is indication here of an additional level. Greek restricts stress in one of the three final positions of the word. In

languages with no restrictions in stress position, however, more intermediate levels might be possible, raising the question of whether stress itself is better viewed as a continuum. These intermediate levels of stress can be attributed to the scope of primary stress-induced strengthening. In this view, stress affects a continuous stretch of speech extending both before and after the stressed syllable. These effects peak either within the stressed syllable or within the first post-stress syllable, with spillover effects being larger than anticipatory ones (see also Katsika & Tsai, 2019). For example, in our data, C gestures show their largest displacement on the syllable that immediately follows the stressed one. Note that peak of displacement effects on the post-stress C gesture challenges, on the one hand, the assumed hierarchy of prominence, in the sense that unstressed gestures present stronger effects than stressed ones. It could, on the other hand, be connected to F0 peak delay (cf. Xu, 1997) and/or alignment of the second component of a bitonal pitch accent to the post-stress syllable (cf. Arvaniti, 1998). This latter hypothesis would work well with the concept of articulatory expansion assisting the incorporation of pitch accents in the coordinative structure of the utterance. This adds to the argument presented in Section 4.2.2 for the usefulness of examining the effects of tonal composition of pitch accents on the phonetic manifestation of prominence.

Effects of prominence extending beyond the stressed syllable have been found elsewhere in the literature, although these mainly concern contrastive accent, and not stress itself (e.g., Cambier-Langeveld & Turk, 1999; Dimitrova & Turk, 2010; Katsika & Tsai, 2019; Sluijter & van Heuven, 1995; Turk & White, 1999; White & Turk, 2010). As a reminder, in the current work, it is paradigmatic comparisons that support the presence of anticipatory and spillover effects of stress. Syntagmatic comparisons, on the other hand, provide partial evidence for the presence of secondary stress (cf. Arvaniti, 1992).

4.4. Interactions between prominence and boundaries

Testing Hypothesis IV, we see that gestures undergo the effects of stress regardless of the position of the word in the phrase, highlighting the kinematic signature of stress. However, as expected, the position of stress within the word affects the timing of the effects of prosodic boundaries, and vice versa. To be specific, our analyses confirmed previous findings reported in Katsika (2016), according to which boundary-related lengthening is initiated earlier the earlier the stress is within the word. This pattern is accompanied by shortening of the gestures consisting the initial syllable of the phrase-final word. This timing relationship between stress and boundary-related lengthening also accounts for the finding that the respective kinematic effects are sequential, not additive, to each other. For example, C gestures at the onset of phrase-final syllables undergo stress-, but not boundary-related lengthening. Importantly, the current study further extends these findings to the dimension of displacement, with boundary-adjacent gestures presenting larger and initial gestures of phrase-final words smaller displacement. We consider the parallel effects on duration and displacement indicative of the articulatory mechanism of prominence, which shall be discussed next.

4.5. A dynamical account of prominence

4.5.1. A mass-spring dynamical model of prominence—Our analyses detect concurrence of duration, position and velocity effects for prominence-marking in Greek, and dependency relations between these three dimensions. However, duration and displacement emerge as dimensions that are also controlled independently, while modification in velocity is fully accounted for by modification in displacement. The implications of this for a mass-spring dynamical model (cf. proposals in Harrington et al., 1995; Cho, 2006) is that there is not a single control parameter (i.e., truncation, rescaling, target, and stiffness) that could account on its own for articulatory control during prominence. This conclusion is in accordance with previous discussions on that matter (e.g., Beckman et al., 1992; Cho, 2006; de Jong, 1995; Harrington et al., 1995; Mücke & Grice, 2014). Either *truncation* or *rescaling* would result in longer and larger, but not faster movements, *target* modification would produce larger and faster, but not longer movements, and finally *stiffness* modification would make the movements longer, larger and slower, instead of faster (see discussion in Cho, 2006; Mücke & Grice, 2014). Possibly, a combination of target modification, which would produce velocity along with displacement effects, with either truncation or rescaling, could capture the patterns observed here.

4.5.2. A μ -gesture model of prominence—The μ -gesture model could provide an alternative account. Within this model, given the patterns found here, lexical stress in Greek would require the concurrent activation of a temporal and a spatial μ -gesture. The temporal μ -gesture would give rise to longer motions of the constrictions gestures coactive with it, while the spatial μ -gesture would change the target parameter for those gestures, strengthening their motion. The faster movements observed in our data would be the outcome of the codependency between amplitude of motion and its peak velocity, according to which the peak velocity value changes in proportion to the change of the target value (e.g., Byrd & Saltzman, 1998; Kelso et al., 1985; Sorensen & Gafos, 2016).

Using the μ -gesture model to account for prominence has some advantageous features. First and foremost, μ -gestures are combinatorial in nature. The μ -model treats the temporal domain separate from the spatial domain, and, similarly to the rate of vibration of the vocal folds (F0), modulations in these domains are considered atomic linguistic units, controllable independently and in parallel to parameters internal to constriction gestures. This makes the theoretical implication that μ -gestures are combinable with other atomic linguistic units, such as constriction gestures, tone gestures, or other μ -gestures, in building larger linguistic constituents. Like all other types of gestures, they are specified for an activation interval, are coordinated with the other gestures that belong to the same coordinative structure with them, and extend their influence over all the gestures that are coactive with them. In this way, they account for effects that extend outside a single constriction gesture or even syllable. Hence, contrary to the mass-spring model discussed above, which needs to be specified for each gesture affected by prominence individually, μ -gestures affect all constriction gestures that overlap with their activation interval.

Following this reasoning, we propose that lexical stress-activating μ -gestures are part of the syllable/word's coordinative structure in the same way that lexical tones are in

a tone language (cf. Gao, 2008). Consequently, the so-called anticipatory (before the stressed syllable) and spillover (after the stressed syllable) effects in fact emerge from this μ -gesture's activation interval (e.g., Cambier-Langeveld & Turk, 1999; Dimitrova & Turk, 2010; Katsika & Tsai, 2019; Sluijter & van Heuven, 1995; Turk & White, 1999; White & Turk, 2010). The fact that anticipatory and spillover effects are less strong than the effects on the stressed syllables is due to the shape of μ -gestures, shown in Fig. 26. μ -Gestures are assumed to be similar to the shape of the π -gestures. This shape accounts for how μ -gestures' effects have a peak, lasting as long as the high plateau of the modulation gesture, and decrease with distance from it (cf. Byrd & Saltzman, 2003). Note that a mass-spring model cannot account directly for anticipatory and spillover effects. For such effects to arise, individual gestures of the stress-adjacent syllables would need to be specified for modifications of lesser degree as compared to those affecting the stressed syllable. Specifying gestures in this way would complicate both planning and computation processes.

The μ -gesture model views prosodic spatial and temporal modulations as separate linguistic units, allowing in this way for typological variation concerning the type of lexical prosody (e.g., languages with stress vs. languages without stress) or the number and/or degree of stress correlates a given language uses (see overview in Fletcher, 2010). For instance, a combination of a temporal μ -gesture with a spatial μ -gesture could capture a stress-language like Greek, while absence of any μ -gesture a language with no lexical stress at all, like, for example, Korean. Similarly, a single temporal μ -gesture could capture a language that uses duration as a stress correlate, while a single spatial gesture a language that uses solely amplitude. The typological power of the μ -gesture model extends to the higher levels of hierarchy of prominence. In particular, languages, in addition to using tone gestures for the appropriate pitch accent, may denote different degrees of phrasal prominence by controlling the strength of μ -gestures. This is similar to how the strength of p -gesture corresponds to the level of prosodic boundary it instantiates (see Fig. 26; cf. Byrd & Saltzman, 2003).

Contrary to accounts in which amplitude and duration are interdependent (e.g., truncation and rescaling), the combinatorial property of μ -gestures is also compatible with the concepts of redundancy and cue trading (e.g., Arvaniti, 1991, 2000; Dauer, 1980). Think of a language that combines a temporal and a spatial μ -gesture to mark prominence. Cue trading may be occurring in cases in which one of the two expected μ -gestures is reduced or deleted due to planning or execution errors or mismatches. In these cases, the other μ -gesture might be enhanced, through relevant feedback and in virtue of redundancy, as a form of compensation.

Last, but not least, μ -gestures work well with the hyperarticulation account (de Jong, 1995), by locally causing more extreme articulatory movements. Modulation gestures are less agreeable with the sonority account (Beckman et al., 1992), since their effect applies to all gestures coactive with them, and does not isolate solely the jaw. Notably, our results are more in accordance with the hyperarticulation account, since both the labial and the lingual systems are found here to present more extreme movements under prominence (cf. discussion in Mücke & Grice, 2014).

4.5.3. Comparing phrasal prominence- to phrasal boundary-instantiating modulation gestures—The current work belongs to a series of studies that in combination could provide an insight into how the mechanisms that instantiate phrasal prominence and boundaries differ. Phrasal prominence and boundaries have been described as prosodic positions with similar articulatory signatures, albeit with distinct – and perceivable as such – functions. Phrasal prominence involves longer, larger and faster gestures, while phrasal boundaries longer, larger but slower gestures (cf. Cho, 2006). The question thus arises what makes phrasal prominence and boundaries kinematically distinct. We propose that both boundaries and prominence have a temporal and a spatial component, and that it is the spatial component that differentiates between the two prosodic functions.

Specifically, in this proposal, the temporal component of both prominence and boundaries is instantiated by temporal μ -gestures, usually called π -gestures when related to boundaries, the function of which is to modulate the rate of utterance time flow. However, the spatial component of boundaries corresponds to pause postures, i.e., specific articulatory configurations marking pauses (Katsika, 2016), while that of prominence to a spatial μ -gesture (cf. Saltzman et al., 2008). In the case of strong boundaries, their spatial component results in audible pauses, whereas less strong boundaries do not achieve the pause posture target. This is because we view boundaries as a continuum corresponding to the continuous activation of π -gestures. When these π -gestures reach a certain threshold at their level of activation, they trigger tone gestures, and at a higher threshold, pause postures are triggered (see Fig. 27(i); for a detailed description of this account see Katsika, 2016).

In this view, pause postures are the spatial target of boundaries, and although all boundaries “move” towards this target, only boundaries that reach the corresponding threshold actually achieve the target. The question then is what prevents less strong boundaries from reaching the threshold required for pause posture activation. One possibility could be truncation by an overlapping π -gesture that initiates the following phrase. In this case a boundary’s strength does not need to be part of planning, but emerges from the timing of π -gestures. When the timing of the following π -gesture is sufficiently later, or if there is no phrase to follow, and thus no other π -gesture to be included in the coordinative structure, a pause posture is reached and a pause is heard.

The spatial component of prominence, on the other hand, is instantiated by a spatial μ -gesture, the function of which is to augment its co-active constriction gestures by smoothly changing their target parameter. Once again, we can view this change as a continuum giving rise to different degrees of prominence, and not as planned, targeted level of activation.

The different velocity profiles of boundaries and prominence thus accrue from their different spatial component. Boundaries “head” towards a pause posture, i.e., a halt of movement and velocity zero. Prominence’s μ -gestures, on the other hand, modify the target of the constriction gestures overlapping with them, which, depending on contextual factors (such as blending strength of overlapped gestures and whether the articulator comes into contact with a hard structure), increase amplitude of motion, and consequently velocity as well. As a reminder, peak velocity changes in proportion to the change in displacement, as found here

and elsewhere in the literature (e.g., Byrd & Saltzman, 1998; Kelso et al., 1985; Sorensen & Gafos, 2016).

It is unclear based on current evidence whether the two μ -gestures, the temporal and the spatial one, involved in marking phrasal prominence are independent from each other or not. One possibility is that each μ -gesture is autonomous and directly coordinated with the constriction gestures of coordinative structure in which they belong (Fig. 27(ii.a)). Another possibility, that is a closer equivalent of the model proposed for the boundaries, is that one μ -gesture is activated when the other μ -gesture reaches a given level of activation (Fig. 27(ii.b)). Which of the two μ -gestures is activated first might be dependent on the type of lexical prosody the language in question uses. For instance, in a language which marks lexical stress primarily by the means of amplitude, it might be the spatial μ -gesture that triggers the temporal one. According to this hypothesis, the spatial μ -gesture would instantiate lexical stress. The strength of that μ -gesture would increase for marking phrasal prominence. When this gesture would reach the appropriate level of activation, the temporal μ -gesture would be triggered, which, as soon as it reached its own adequate threshold, would in turn trigger the relevant pitch accent (see Fig. 27(ii.b)). Future work will need to test the validity of the models of phrasal prominence and boundaries proposed here and to examine their different alternative versions by the means of both cross-linguistic empirical data and computational modeling data.

Finally, to capture the timing of boundary-related events in Greek, our previous work has proposed a dual coordination relationship for the π -gesture that instantiates the boundary, meaning a) with the final V gesture, and b) with the temporal μ -gesture that instantiates the stress of the phrase-final word (Katsika, 2016; see Fig. 27(ii)). In that proposal, the coordination with the V gesture was characterized as anti-phase (sequential), but the coordination with the stress-related μ -gesture was left uncharacterized. In light of the current findings of boundary- and stress-related lengthening not being additive, we can also assume a sequential, anti-phase, coordination between the π - and the μ -gestures, as shown in Fig. 27(i). The interaction of the two coordination relationships is such that, when stress is not final, the stress-related μ -gesture is slightly pulled towards the right word boundary, accounting hence for the attenuating effects on the constriction gestures (shorter durations and smaller in displacement) detected at the beginning of phrase-final words.

5. Conclusion

Stress in Greek makes constriction gestures longer and larger. Gestures also become faster the larger they become. These modulations are primarily located in the stressed syllable, and have strong spillover and weaker anticipatory effects. Accent does not enhance these modulations, although it exerts kinematic effects of minimal magnitude across the whole word. With the exception of duration of final V gestures, these effects do not encode focus type. Based on these results, we put forward a gestural account, according to which prominence in Greek is instantiated by the concurrence of two μ -gestures, a temporal and a spatial one, activated at the lexical stress level. At phrase edges, prominence-inducing μ -gestures are coordinated with boundary-inducing ones, giving rise to the intricate set of events that mark prosodic boundaries. Other languages may differ in the number and type

of μ -gestures used for inducing prominence, the level of prominence at which these gestures are activated, and how they are coordinated, if at all, with boundary-inducing μ -gestures.

Acknowledgments

This work was supported by the National Science Foundation [# 1551428] and National Institutes of Health [DC 002717 and DC 008780].

References

- Arvaniti A (1991). Rhythmic categories: a critical evaluation on the basis of Greek data. In Proceedings of the XIIth international congress of phonetic sciences (pp. 298–301). Université de Provence, Service des Publications.
- Arvaniti A (1992). Secondary stress: Evidence from Modern Greek. In Docherty GJ & Ladd DR (Eds.), *Papers in laboratory phonology II: Gesture, Segment, Prosody* (pp. 398–423). Cambridge University Press.
- Arvaniti A (1998). Phrase accents revisited: Comparative evidence from Standard and Cypriot Greek. In Proceedings of the 15th international congress of phonetic sciences (pp. 2883–2886).
- Arvaniti A (2000). The acoustics of stress in Modern Greek. *Journal of Greek Linguistics*, 1, 9–39.
- Arvaniti A (2007). Greek phonetics: The state of the art. *Journal of Greek Linguistics*, 8, 97–208.
- Arvaniti A, & Baltazani M (2005). Intonational analysis and prosodic annotation of Greek spoken corpora. In Jun S-A (Ed.), *Prosodic typology: The phonology of intonation and phrasing* (pp. 84–117). Oxford: Oxford University Press.
- Arvaniti A, & Ladd DR (2009). Greek wh-questions and the phonology of intonation. *Phonology*, 26, 43–74.
- Arvaniti A, Ladd DR, & Mennen I (2006). Tonal association and tonal alignment: Evidence from Greek polar questions and contrastive statements. *Language and Speech*, 49, 421–450. [PubMed: 17326587]
- Astruc L, & Prieto P (2006). Acoustic cues of stress and accent in Catalan. In Hoffmann R & Mixdorff H (Eds.), *Proceedings of speech prosody 2006* (pp. 337–40). Dresden: TUD Press.
- Avesani C, Vayra M, & Zmarich C (2007). On the articulatory basis of prominence in Italian. In Proceedings of the XVIth international congress of phonetic sciences (pp. 981–984). Saarbrücken, Germany.
- Baltazani M (2006). Intonation interpretation of negation in Greek. *Journal of Pragmatics*, 38, 1658–1676.
- Baltazani M (2007). Prosodic rhythm and the status of vowel reduction in Greek. In *Selected papers on theoretical and applied linguistics from the 17th international symposium on theoretical & applied linguistics* (vol. 1, pp. 31–43). Department of Theoretical and Applied Linguistics, Thessaloniki.
- Baltazani M, & Jun S-A (1999). Focus and topic intonation in Greek. In Proceedings of the XIVth international congress of phonetic sciences (pp. 1305–1308). San Francisco, USA.
- Baumann S, Grice M, & Steindamm S (2006). Prosodic marking of focus domains – Categorical or gradient? In Proceedings of speech prosody 2006 (pp. 301–304). Dresden, Germany.
- Beckman ME, & Edwards J (1992). Intonational categories and the articulatory control of duration. In Perception S (Ed.), *Production and linguistics structure* (pp. 359–375). Tokyo, Japan: Ohmsha.
- Beckman ME, Edwards J, & Fletcher J (1992). Prosodic structure and tempo in a sonority model of articulatory dynamics. In Docherty GJ & Ladd DR (Eds.), *Papers in laboratory phonology II: Segment, gesture, prosody* (pp. 68–86). Cambridge: Cambridge University Press.
- Beckman ME, & Edwards J (1994). Articulatory evidence for differentiating stress categories. In Kearing PA (Ed.), *Phonological structure and phonetic Evidence* (pp. 7–33). Cambridge: Cambridge University Press.
- Beckman ME, & Pierrehumbert J (1986). Intonation structure in English and Japanese. *Phonology Yearbook*, 3, 255–310.

- Bombien L, Mooshammer C, Hoole P, Rathcke T, & Kuehnert B (2007). Articulatory strengthening in initial German /kl/ clusters under prosodic variation. In Triuvain J & Barry WJ (Eds.), *Proceedings of the XVIth international congress of phonetic sciences* (pp. 457–60), Saarbücken, Germany.
- Botinis A (1982). *Stress in modern Greek: An acoustic study*. Working Papers Linguistics-Phonetics Lund University, 22, 27–38.
- Botinis A (1989). *Stress and prosodic structure in Greek: A phonological, acoustic, physiological and perceptual study* Lund: Lund University Press.
- Botinis A (1998). Greek intonation. In Hirst D & Di Cristo A (Eds.), *Intonation systems: A survey of twenty languages* (pp. 280–310). Cambridge: Cambridge University Press.
- Botinis A, Fourakis M, & Katsaiti M (1995). Acoustic characteristics of Greek vowels under different prosodic conditions In *Proceedings of the XIIIth international congress of phonetic sciences* (pp. 404–407). Stockholm: KTH and Stockholm University.
- Botinis A, Fourakis M, Panagiotopoulou N, & Pouli K (2001a). Greek vowel durations and prosodic interactions. *Glossologia*, 13, 101–123.
- Botinis A, Fourakis M, & Bannert R (2001). Prosodic interactions on segmental durations in Greek In *Proceedings of the XIVth Swedish phonetics conference PNETIK 2001 - Lund University, Dept. of Linguistics Working Papers*, 49, 10–13.
- Botinis A, & Bannert R (2003). Focus and gender interactions and prosodic correlates. *Phonum*, 9, 105–108.
- Botinis A, Bannert R, & Tzimokas D (2002). Cross-linguistic prosody of emphasis among Men and Women] *Proceedings of the 6th international conference on Greek linguistics*.
- Browman CP, & Goldstein LM (1986). Towards an articulatory phonology. *Phonology Yearbook*, 3, 219–252.
- Browman CP, & Goldstein LM (1992). Articulatory phonology: An overview. *Phonetica*, 45, 155–180.
- Byrd D, & Saltzman E (1998). Intra-gestural dynamics of multiple phrasal boundaries. *Journal of Phonetics*, 26, 173–199.
- Byrd D, & Saltzman E (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, 31, 149–180.
- Byrd D, Kaun A, Narayanan S, & Saltzman E (2000). Phrasal signatures in articulation. In Kingston J & Beckman ME (Eds.), *Papers in laboratory phonology V: Acquisition and the Lexicon* (pp. 70–87). Cambridge, U.K.: Cambridge University Press.
- Cambier-Langeveld T, & Turk AE (1999). A cross-linguistic study of accentual lengthening: Dutch vs. English. *Journal of Phonetics*, 27, 255–280.
- Cho T (2002). *The effects of prosody on articulation in English* New York: Routledge.
- Cho T (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, 32, 141–176.
- Cho T (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /a, i/ in English. *Journal of the Acoustical Society of America*, 117(6), 3867–3878. [PubMed: 16018489]
- Cho T (2006). Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English. In Goldstein L, Whalen DH, & Best CT (Eds.), *Papers in laboratory phonology VIII: Varieties of phonological competence (phonology and phonetics)* (pp. 519–548). Berlin: Mouton de Gruyter.
- Cho T, & Keating P (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, 37, 466–485.
- Cho T, Kim J, & Kim S (2013). Preboundary lengthening and preaccentual shortening across syllables in a trisyllabic word in English. *Journal of the Acoustical Society of America*, 133(5), EL384–EL390. [PubMed: 23656098]
- Cho T, & McQueen JM (2005). Prosodic influences on consonant production in Dutch: Effects of prosodic boundaries, phrasal accent and lexical stress. *Journal of Phonetics*, 33(2), 121–157.
- Crosswhite K (2003). Spectral tilt as a cue to word stress in Polish, Macedonian, and Bulgarian. In Solé MJ, Recasens D & Romero J (Eds.), *Proceedings of the XVth international congress of the phonetic sciences* (pp. 767–70). Barcelona/Australia: Causal Productions.

- Crystal TH, & House AS (1988). Segmental durations in connected speech signals: Syllabic stress. *Journal of the Acoustic Society of America*, 83, 1574–1585.
- Dauer RM (1980). Stress and rhythm in modern Greek Ph.D. dissertation University of Edinburgh.
- de Jong K (1991). An articulatory study of consonant-induced vowel duration changes in English. *Phonetica*, 48, 1–17. [PubMed: 1745665]
- de Jong KJ (1995). The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticulation. *Journal of the Acoustical Society of America*, 97(1), 491–504. [PubMed: 7860828]
- de Jong K (2004). Stress, lexical focus, and segmental focus in English: Patterns of variation in vowel duration. *Journal of Phonetics*, 32(4), 493–516.
- de Jong K, Beckman ME, & Edwards J (1993). The interplay between prosodic structure and coarticulation. *Language and Speech*, 36(2–3), 197–212. [PubMed: 8277808]
- de Jong K, & Zawaydeh B (2002). Comparing stress, lexical focus, and segmental focus: Patterns of variation in Arabic vowel duration. *Journal of Phonetics*, 30, 53–75.
- Dimitrova S, & Turk AE (2010). Patterns of accentual lengthening in English four-syllable words. *Journal of Phonetics*, 40, 403–418.
- Dohen M, & Loevenbruck H (2005). Audiovisual production and perception of contrastive focus in French: A multispeaker study. *Interspeech*, 2005, 2413–2416.
- Dohen M, Loevenbruck H, & Hill H (2006). Visual correlates of prosodic contrastive focus in French: Description and inter-speaker variabilities In *Proceedings of Speech Prosody 2006* (pp. 221–224). Dresden, Germany.
- Dogil G (1999). The phonetic manifestation of word stress in Lithuanian, Polish, German, and Spanish. In van der Hulst H (Ed.), *Word prosodic systems of the languages of Europe* (pp. 272–310). Berlin: Mouton de Gruyter.
- Edwards J, Beckman ME, & Fletcher J (1991). The articulatory kinematics of final lengthening. *Journal of the Acoustical Society of America*, 89(1), 369–382. [PubMed: 2002175]
- Fant G, Kruckenberg A, & Nord L (1991). Durational correlates of stress in Swedish, French and English. *Journal of Phonetics*, 19, 351–365.
- Fletcher J (2010). The prosody of speech: Timing and rhythm. In Hardcastle WJ, Laver J & Gibbon FE, (Eds.), *The handbook of phonetic sciences* (pp. 523–602). United Kingdom: Wiley-Blackwell Publishing.
- Fourakis M (1991). Tempo, stress, and vowel reduction in American English. *Journal of the Acoustical Society of America*, 90, 1816–1827. [PubMed: 1960277]
- Fourakis M, Botinis A, & Katsaiti M (1999). Acoustic characteristics of Greek vowels, *Phonetica* 56.28–43. [PubMed: 10450074]
- Fowler CA (1995). Acoustic and kinematic correlates of contrastive stress accent in spoken English. In Bell-Berti F & Raphael JJ (Eds.), *Producing speech: Contemporary issues: For Katherine Safford Harris* (pp. 355–373). Woodbury: American Institute of Physics.
- Goldstein LM, Byrd D, & Saltzman E (2006). The role of vocal tract gestural action units in understanding the evolution of phonology. In Arbib M (Ed.), *From action to language: The mirror neuron system* (pp. 215–249). Cambridge: Cambridge University Press.
- Gussenhoven C (2004). *The phonology of tone and intonation* Cambridge: Cambridge University Press.
- Gussenhoven C (2009). Vowel quantity, syllable duration, and stress in Dutch. In Hanson K & Inkelas S (Eds.), *The nature of word: Essays in honor of Paul Kiparsky* (pp. 181–198). Cambridge: Cambridge University Press.
- Harrington J, Fletcher J, & Beckman ME (2000). Manner and place conflicts in the articulation of accent in Australian English. In Broe M (Ed.), *Papers in laboratory phonology 5* (pp. 40–55). Cambridge: Cambridge University Press.
- Harrington J, Fletcher J, & Roberts C (1995). Coarticulation and the accented/unaccented distinction: Evidence from jaw movement data. *Journal of Phonetics*, 23 (3), 305–322.

- Hawkins S (1992). An introduction to task dynamics. In Docherty J & Ladd DR (Eds.), *Papers in laboratory phonology II: Gesture, segment, prosody* (pp. 9–25). Cambridge: Cambridge University Press.
- Hayes B (1989). The prosodic hierarchy in meter. In Kiparsky P & Youmans G (Eds.), *Rhythm and meter* (pp. 201–260). Orlando, FL: Academic Press.
- Heldner M, & Strangert E (2001). Temporal effects of focus in Swedish. *Journal of Phonetics*, 29, 329–361.
- Hermes A, Becker J, Mücke D, Baumann S, & Grice M (2008). Articulatory gestures and focus marking in German In *Proceedings of speech prosody 2008* (pp. 457–460). Campinas, Brazil.
- Hoole P, Zierdt A, & Geng C (2003). Beyond 2D in articulatory data acquisition and analysis In *Proceedings of the XVth international congress of phonetic sciences* (pp. 265–268).
- Kastrinaki A (2003). The Temporal correlates of lexical and phrasal stress in Greek, exploring rhythmic stress: Durational patterns for the case of Greek words M.Sc. Dissertation University of Edinburgh.
- Katsika A (2016). The role of prominence in determining the scope of boundary-related lengthening in Greek. *Journal of Phonetics*, 55, 148–181.
- Katsika A, Krivokapic J, Mooshammer C, Tiede M, & Goldstein L (2014). The coordination of boundary tones and their interaction with prominence. *Journal of Phonetics*, 44, 62–82. [PubMed: 25300341]
- Katsika A, & Tsai K (2019). The scope of prominence-induced lengthening in Greek In *Proceedings of the international congress of phonetic sciences 2019*.
- Kelso JAS, Vatikiotis-Bateson E, Saltzman EL, & Kay B (1985). A qualitative dynamic analysis of reiterant speech production: Phase portraits, kinematics and dynamic modelling. *Journal of the Acoustical Society of America*, 77, 266–280. [PubMed: 3973219]
- Kim S, Jang J, & Cho T (2017). Articulatory characteristics of preboundary lengthening in interaction with prominence on tri-syllabic words in American English. *The Journal of the Acoustical Society of America*, 142(4), EL362–EL368. [PubMed: 29092557]
- Kohler K (1983). Prosodic boundary signals in German. *Phonetica*, 40, 89–134.
- Ladd DR, & Morton R (1997). The perception of intonational emphasis: Continuous or categorical? *Journal of Phonetics*, 25, 313–342.
- Lieberman M, & Prince A (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8, 249–336.
- Lucero J, Munhall K, Gracco V, & Ramsay J (1997). On the registration of time and the patterning of speech movements. *JSLHR*, 40, 1111–1117. [PubMed: 9328881]
- Mücke D, & Grice M (2014). The effect of focus marking on supralaryngeal articulation – Is it mediated by accentuation? *Journal of Phonetics*, 44, 47–61.
- Munhall KG, Ostry DJ, & Parush A (1985). Characteristics of velocity profiles of speech movements. *Journal of Experimental Psychology: Human Perception and Performance*, 11, 457–474. [PubMed: 3161986]
- Nam H (2007). Syllable-level intergestural timing model: Split-gesture dynamics focusing on positional asymmetry and moraic structure. *Laboratory Phonology*, 9, 483–506.
- Nespor M, & Vogel I (1986). *Prosodic phonology* Dordrecht: Foris.
- Nicolaidis K (2003). Acoustic variability of vowels in Greek spontaneous speech In *Proceedings of the XVth international congress of phonetic sciences*, 3221–3224. Universidad Autónoma de Barcelona.
- Nicolaidis K, & Rispoli R (2005). The effect of noise on speech production: An acoustic study. *Studies in Greek Linguistics*, 25, 415–426.
- Niebuhr O (2010). On the phonetics of intensifying emphasis in German. *Phonetica*, 67, 170–198. [PubMed: 20926915]
- Nowak PM (2006). *Vowel reduction in Polish* Ph.D. Dissertation Berkeley: University of California.
- Ogden R (2012). Making sense of outliers. *Phonetica*, 69, 48–67. [PubMed: 23172239]
- Oller DK (1979). Syllable timing in Spanish, English and Finnish. In Hollien H & Hollien P (Eds.), *Current issues in the phonetic sciences* (pp. 320–341). Amsterdam: John Benjamins.

- Ostry DJ, & Munhall KG (1985). Control of rate and duration of speech movements. *The Journal of the Acoustical Society of America*, 77, 640–648. [PubMed: 3882804]
- Padgett J, & Tabain M (2005). Adaptive dispersion theory and phonological vowel reduction in Russian. *Phonetica*, 62, 14–54. [PubMed: 16116302]
- Pierrehumbert JB (1980). The phonology and phonetics of English intonation. Ph.D. Dissertation M.I.T
- Pierrehumbert J, & Beckman ME (1988). *Japanese tone structure* Cambridge, MA: M.I.T. Press.
- Prieto P (2015). Intonational meaning. *WIREs Cognitive Science*, 6, 371–381. [PubMed: 26263426]
- R Core Team. (2019). R: A language and environment for statistical computing <https://www.R-project.org/>.
- Rietveld T, Kerkhoff J, & Gussenhoven C (2004). Word prosodic structure and vowel duration in Dutch. *Journal of Phonetics*, 32, 349–371.
- Roon K, Gafos A, Hoole P, & Zeroual C (2007). Influence of articulator and manner on stiffness. In: Trouvain J & Barry W (Eds.), *Proceedings of the 16th international congress of phonetic sciences August 2007* (pp. 409–412). Germany: Saarbrücken.
- Saltzman E, & Munhall KG (1989). A dynamical approach to gestural patterning in speech production. *Ecological Psychology*, 1, 333–382.
- Saltzman E, Nam H, Krivokapi J, & Goldstein L (2008). A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. *Proceedings of the Speech Prosody*, 175–184.
- Selkirk E (1984). *Phonology and syntax: The relation between sound and structure* Cambridge, MA: M.I.T. Press.
- Silverman K, Beckman ME, Pitrelli J, Ostendorf M, Wightman C, Price P, ... Hirschberg J (1992). ToBI: A standard labeling English prosody. *Proceedings of the International Conference on Spoken Language Processing*, 2, 867–870.
- Shattuck-Hufnagel S, & Turk A (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25, 193–247. [PubMed: 8667297]
- Sluijter AMC, & van Heuven VJ (1995). Effects of focus distribution, pitch accent and lexical stress on the temporal organization of syllables in Dutch. *Phonetica*, 52, 71–89.
- Sorensen T, & Gafos A (2016). The gesture as an autonomous nonlinear dynamical system. *Ecological Psychology*, 28, 188–215.
- Turk AE, & Sawusch JR (1997). The domain of accentual lengthening in American English. *Journal of Phonetics*, 25, 25–41.
- Turk AE, & White L (1999). Structural influences on accentual lengthening in English. *Journal of Phonetics*, 27, 171–206.
- Wightman CW, Shattuck-Hufnagel S, Ostendorf M, & Price PJ (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America*, 91, 1707–1717. [PubMed: 1564206]
- Xu Y (1997). Contextual tonal variation in Mandarin. *Journal of Phonetics*, 25, 61–83.

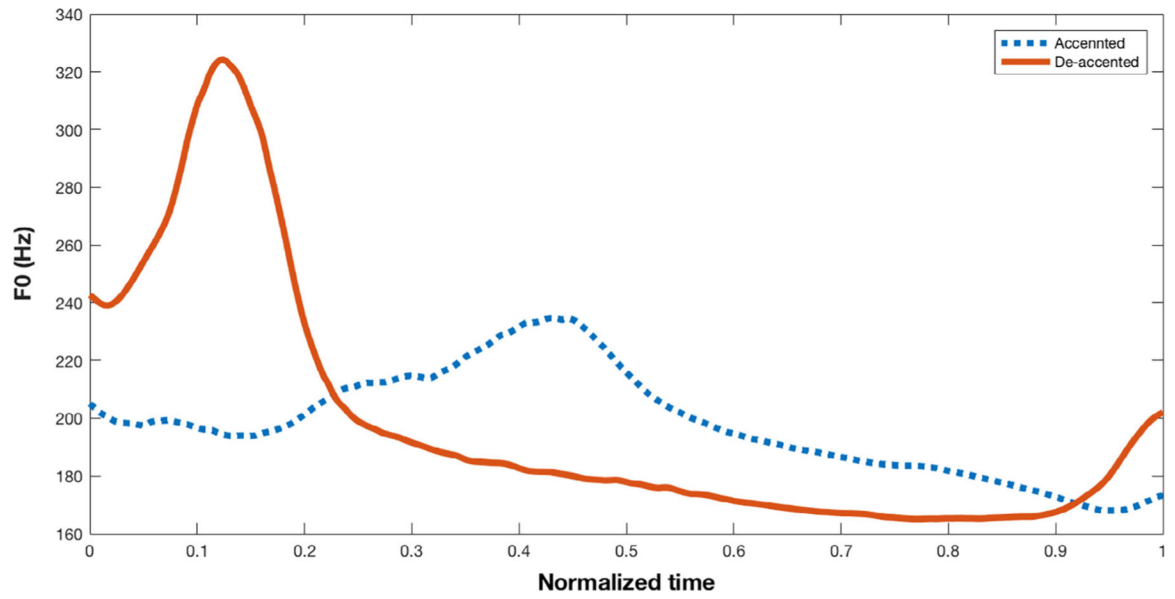


Fig. 1. The mean time-aligned F0 signals of the utterances involving phrase-final test words ($n = 1029$), averaged by accent type (accented vs. de-accented).

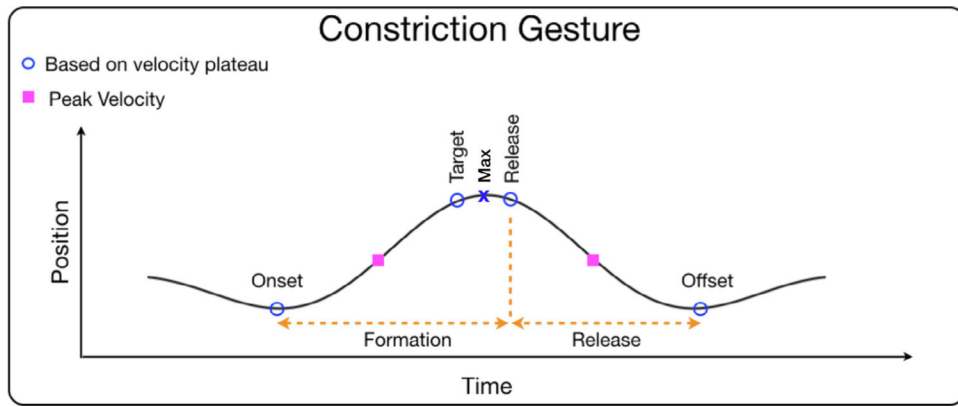


Fig. 2. Constriction gesture's timepoints labeled on the basis of velocity criteria: Peak velocity for the homonymous timepoints, velocity minima for constriction maxima, and velocity plateaus for the other timepoints.

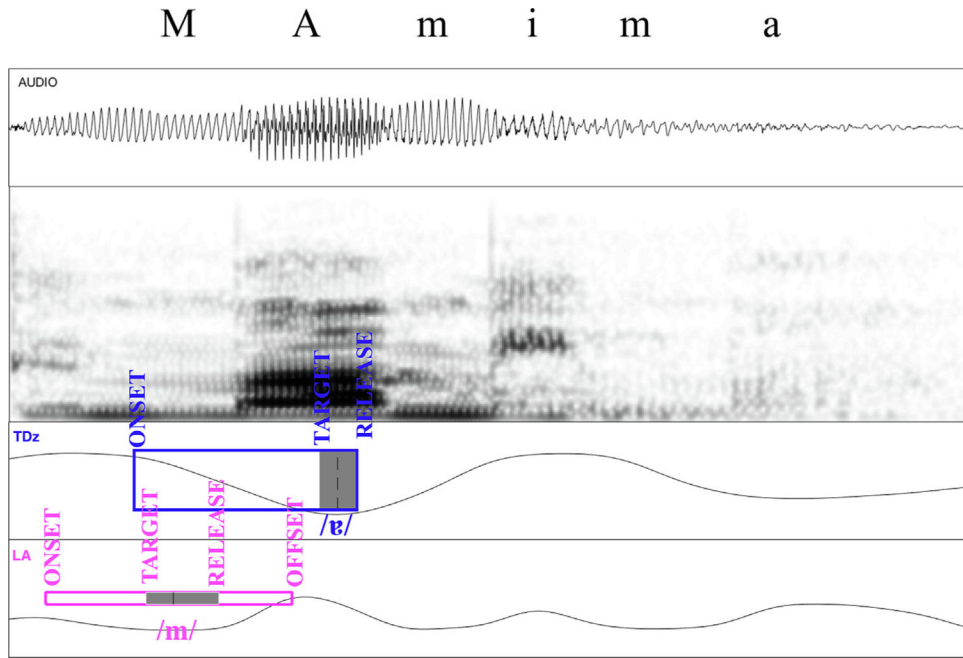


Fig. 3.

The waveform, spectrogram, tongue dorsum vertical displacement (TDz) and lip aperture (LA) of an accented stress-initial word (MAmima) in phrase-final position (IP) produced by speaker F01. The following kinematic labels for the C (/m/) and V (/ə/) of the word's initial syllable are shown on LA (magenta box) and TDz (blue box) respectively: onset, target and release. Solid gray boxes correspond to gestural targets. The vertical broken black line within the solid gray box marks maximal constriction. For C, the offset of the constriction's release is also marked.

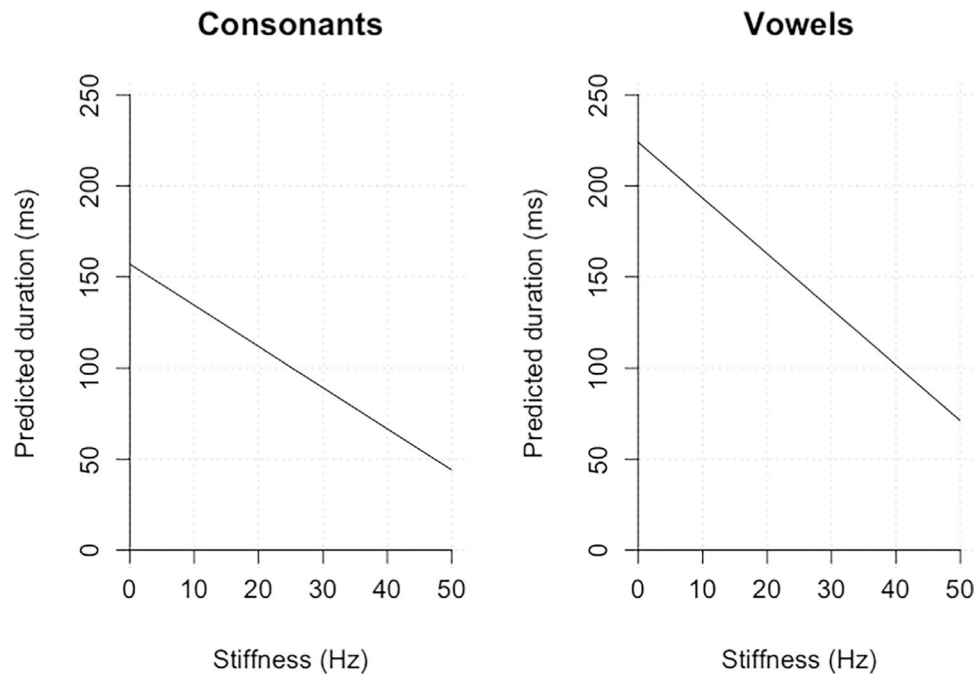


Fig. 4. Relationship between predicted duration (in ms) and stiffness (ratio of peak velocity over displacement; in Hz), based on the β estimate for stiffness, for (a) C gestures and (b) V gestures.

C Gestures

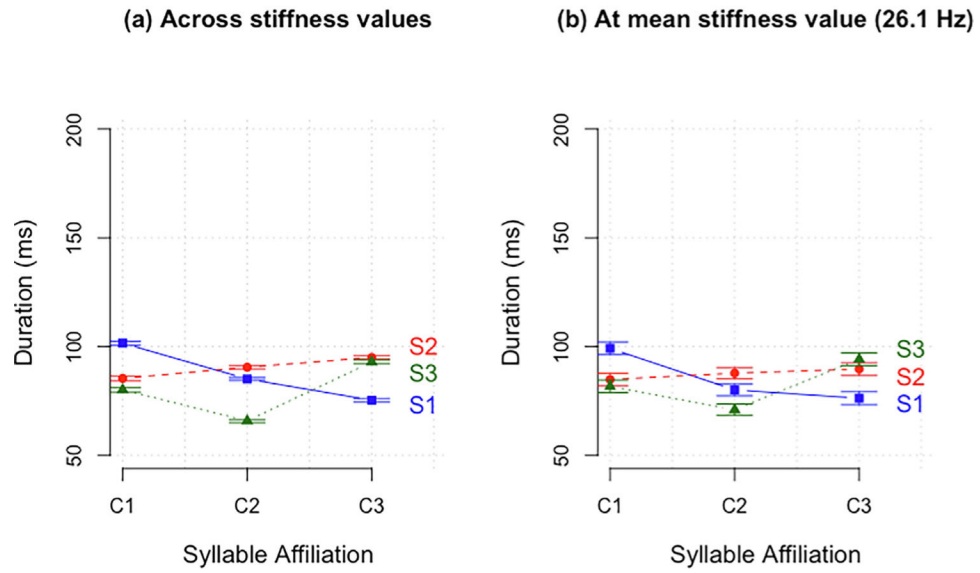


Fig. 5. Formation duration (in ms) of C gestures by syllable affiliation (C1, C2, C3) and stress position (S1, S2, S3) across stiffness values in (a) and at mean value of stiffness (26.1 Hz) in (b).

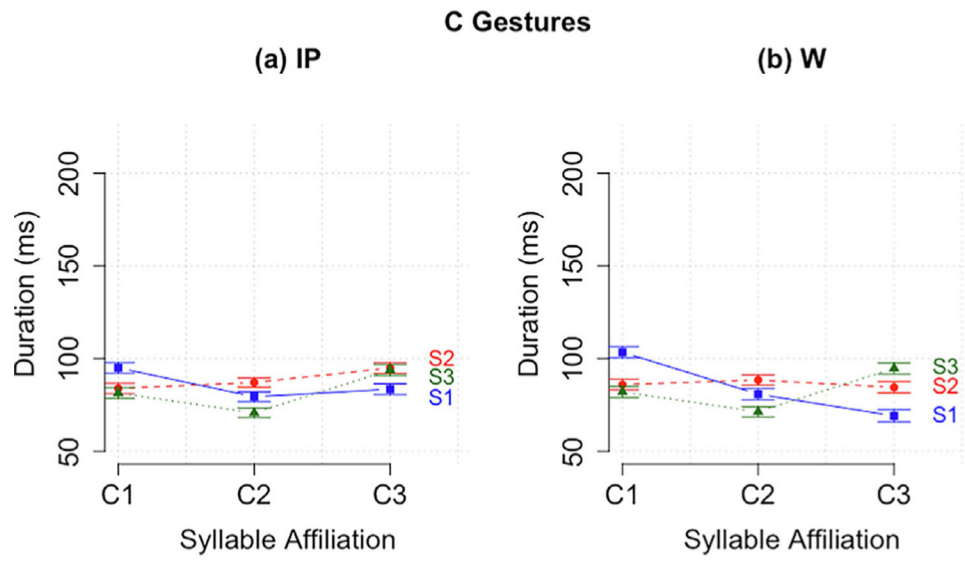


Fig. 6. Formation duration (in ms) of C gestures by syllable affiliation (C1, C2, C3) and stress position (S1, S2, S3) per boundary type (Word and IP).

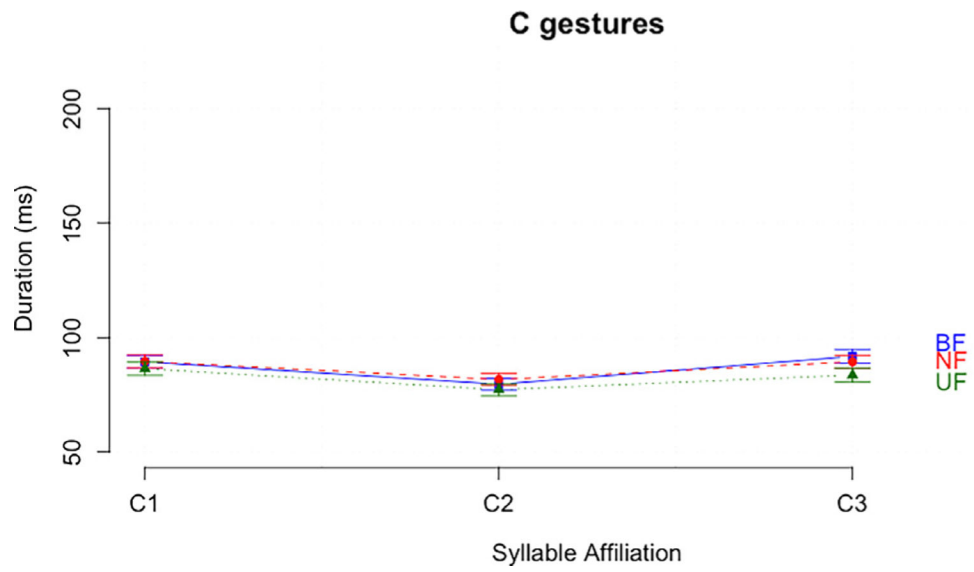


Fig. 7. Formation duration (in ms) of C gestures by syllable affiliation (C1, C2, C3) and focus type (BF, NF, UF).

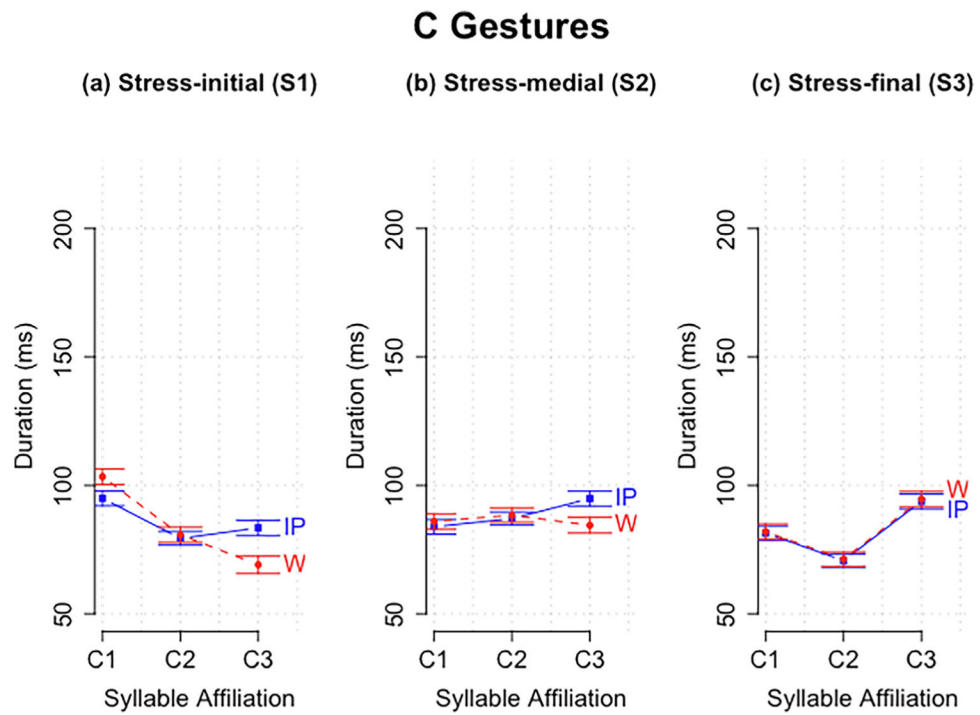


Fig. 8. Formation duration (in ms) of C gestures by syllable affiliation (C1, C2, C3) and boundary type (IP, W) per stress position (S1, S2, S3).

V Gestures

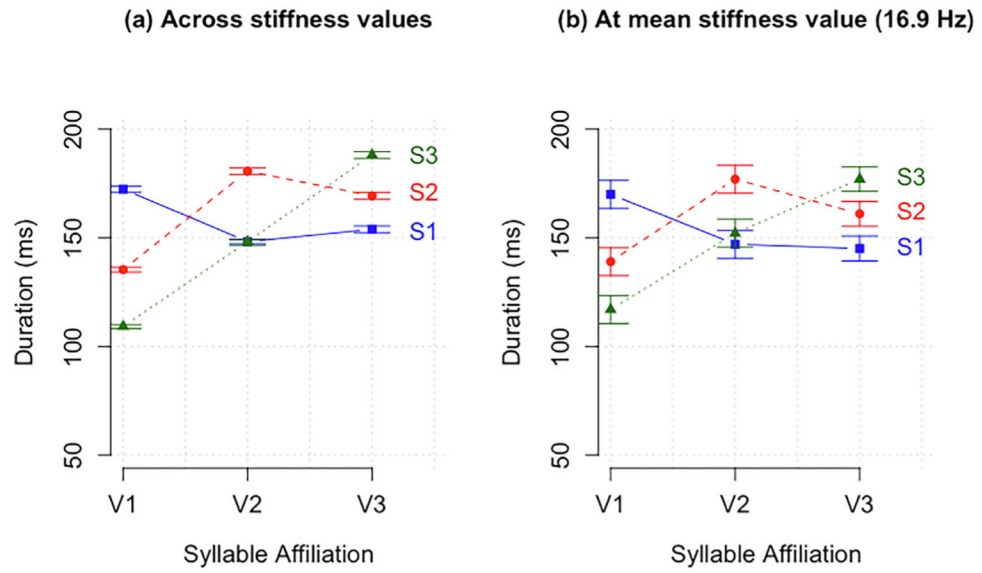


Fig. 9. Formation duration (in ms) of V gestures by syllable affiliation (V1, V2, V3) and stress position (S1, S2, S3) across stiffness values in (a) and at mean value of stiffness (16.9 Hz) in (b).

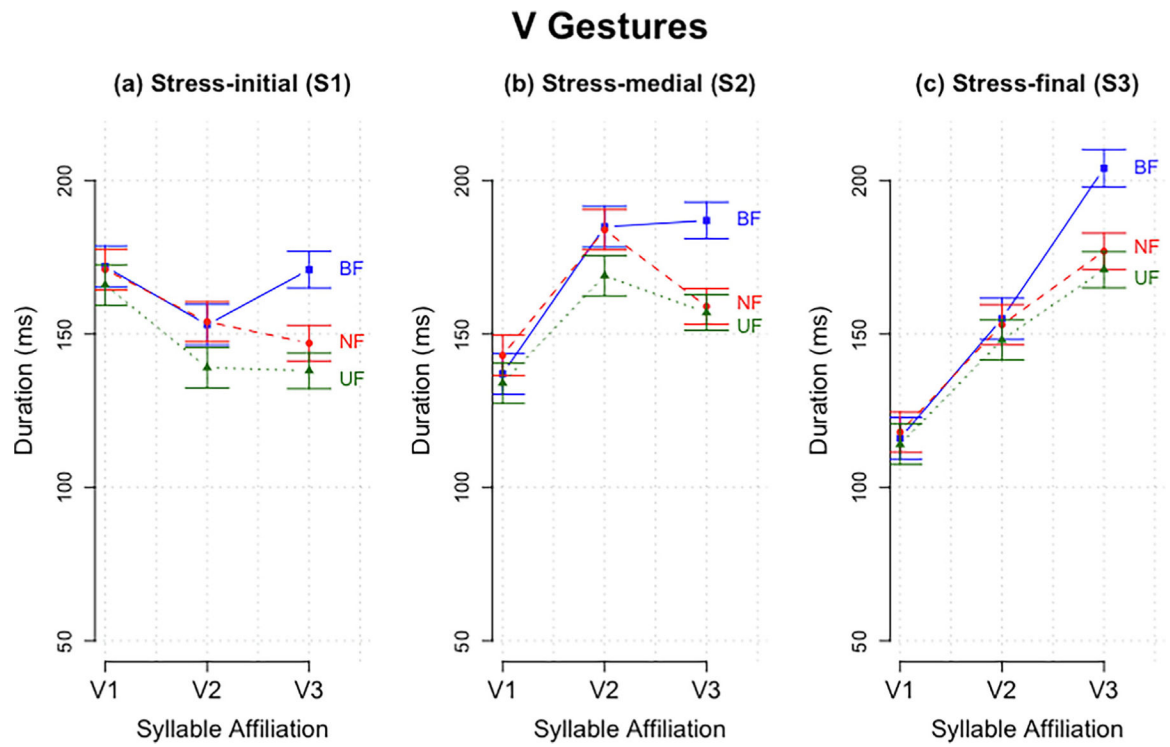


Fig. 10. Formation duration (in ms) of V gestures by syllable affiliation (V1, V2, V3) and focus type position (BF, NF, UF).

V Gestures

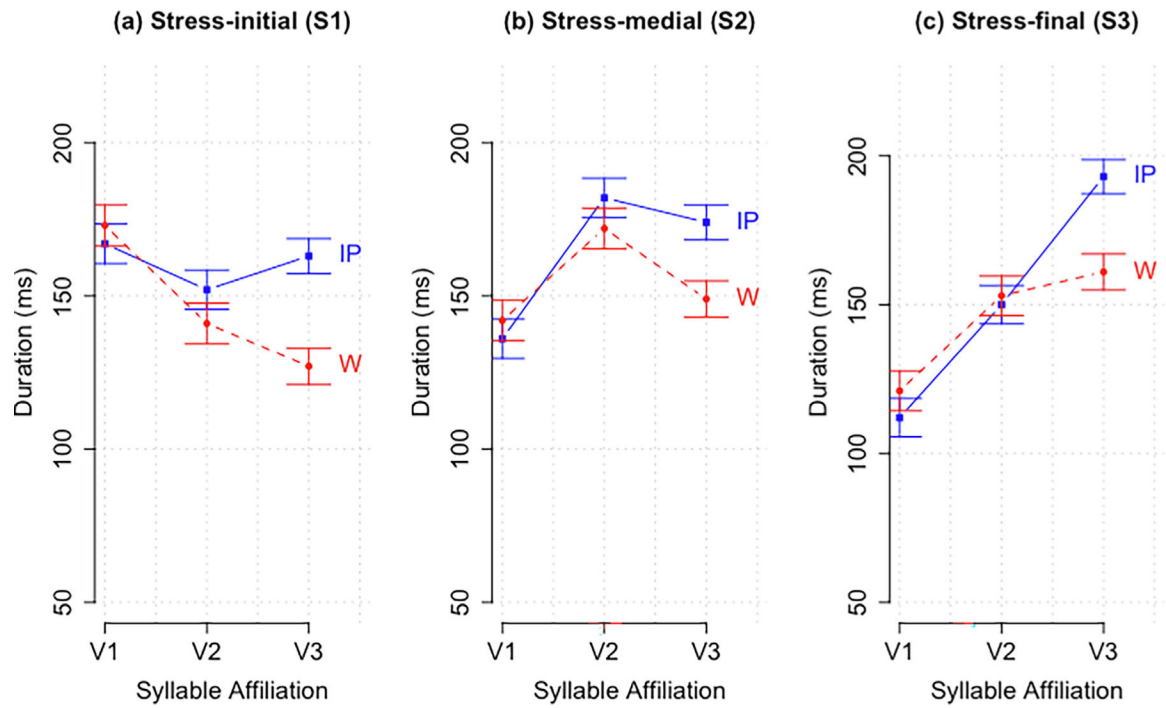


Fig. 11.

Formation duration (in ms) of V gestures by syllable affiliation (V1, V2, V3) and boundary type (IP, W) across stiffness values per stress position (S1, S2, S3).

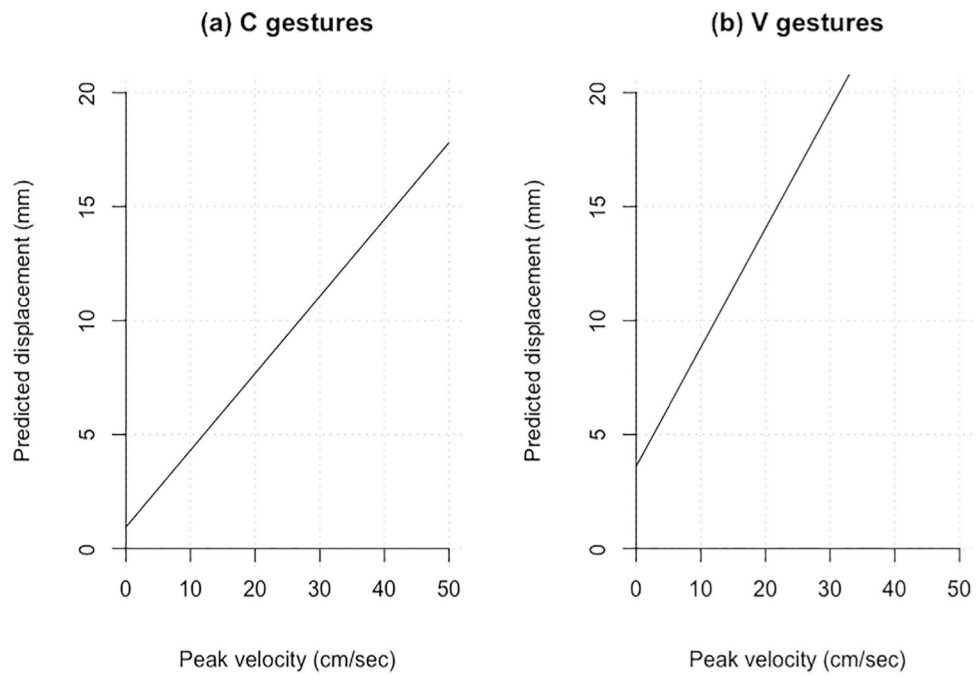


Fig. 12. Relationship between predicted displacement (in mm) and peak velocity (in cm/sec), based on the β estimate of peak velocity, for (a) C gestures and (b) V gestures.

C Gestures

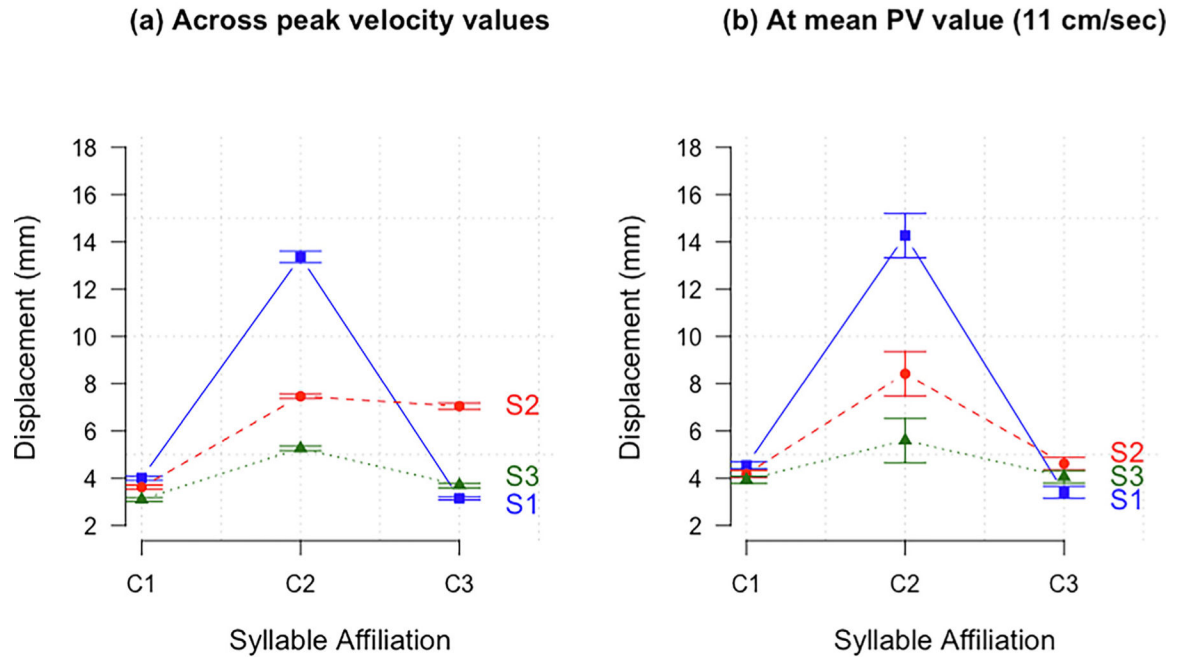


Fig. 13. Displacement (in mm) of C gestures by syllable affiliation (C1, C2, C3) and stress position (S1, S2, S3) across peak velocity values in (a) and at mean value of peak velocity (11 cm/sec) in (b).

C Gestures

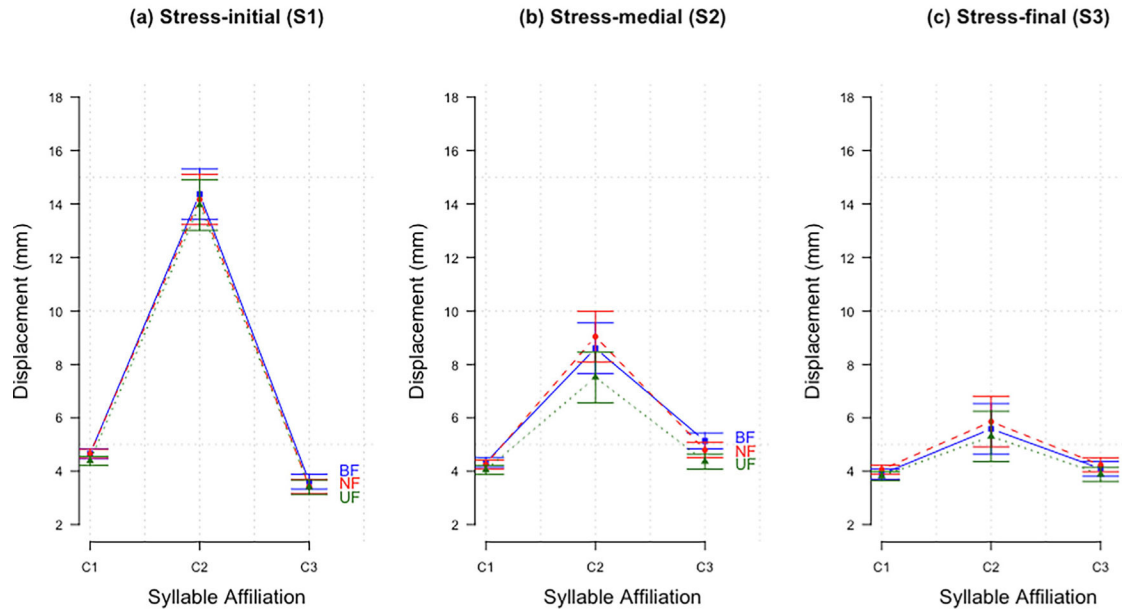


Fig. 14. Displacement (in mm) of C gestures by syllable affiliation (C1, C2, C3) and focus type (BF, NF, UF) across peak velocity values per stress position (S1, S2, S3).

C Gestures

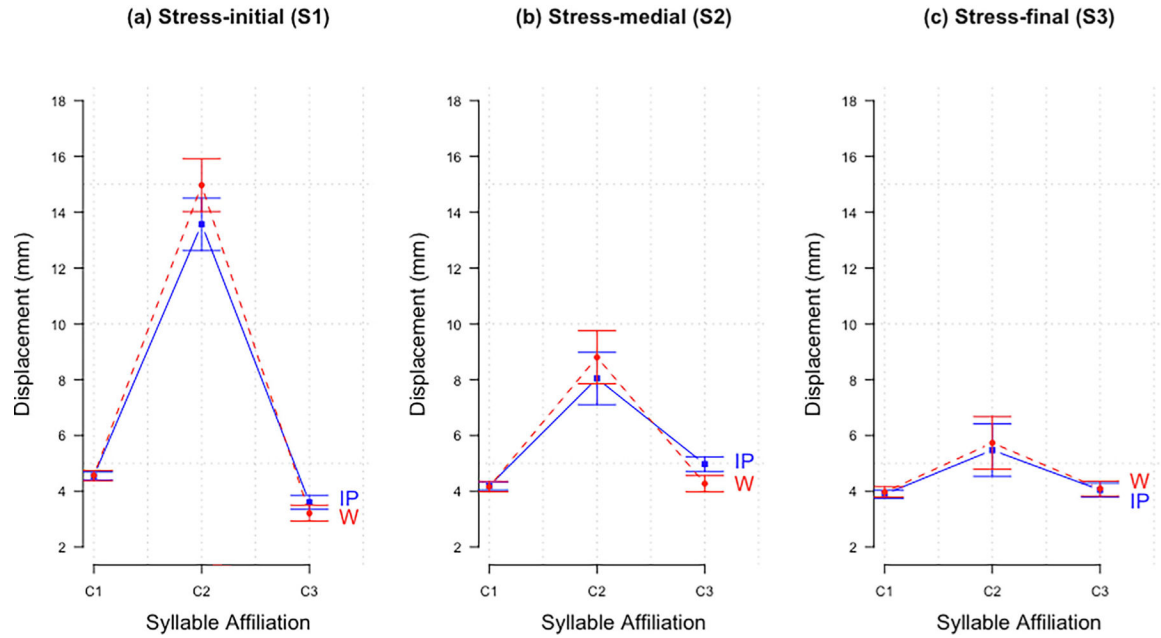


Fig. 15. Displacement (in mm) of C gestures by syllable affiliation (C1, C2, C3) and boundary type (IP, W) across peak velocity values per stress position (S1, S2, S3).

V Gestures

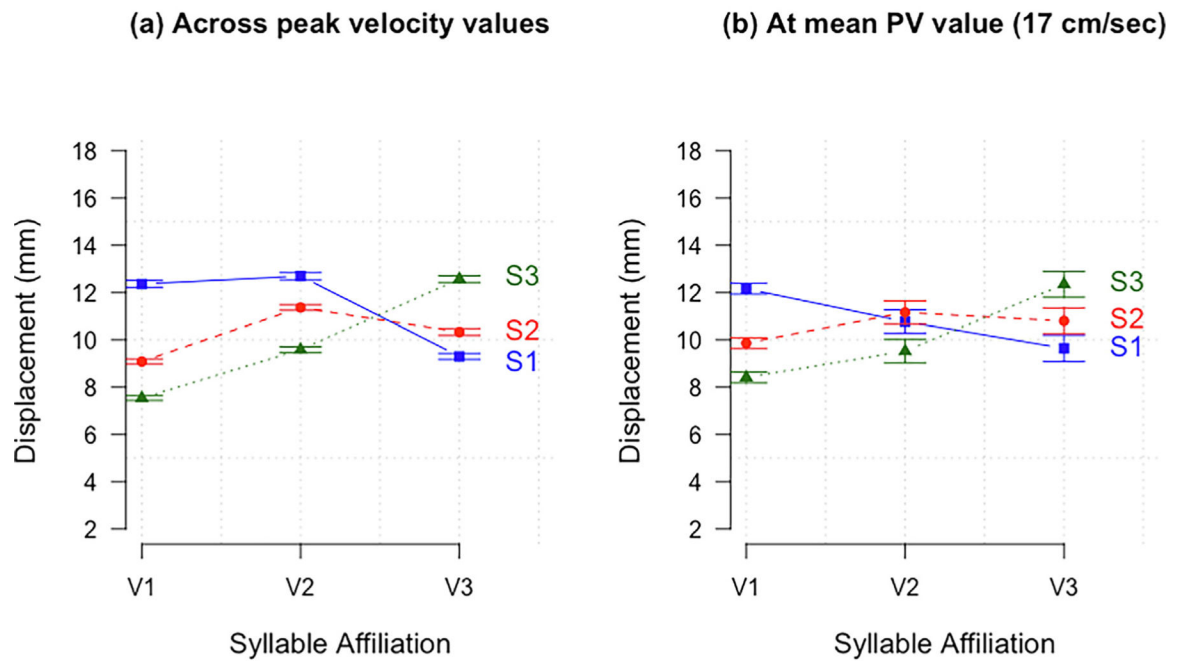


Fig. 16. Displacement (in mm) of V gestures by syllable affiliation (V1, V2, V3) and stress position (S1, S2, S3) across peak velocity values in (a) and at mean value of peak velocity ($mean_{pv} = 17$ cm/sec) in (b).

V Gestures

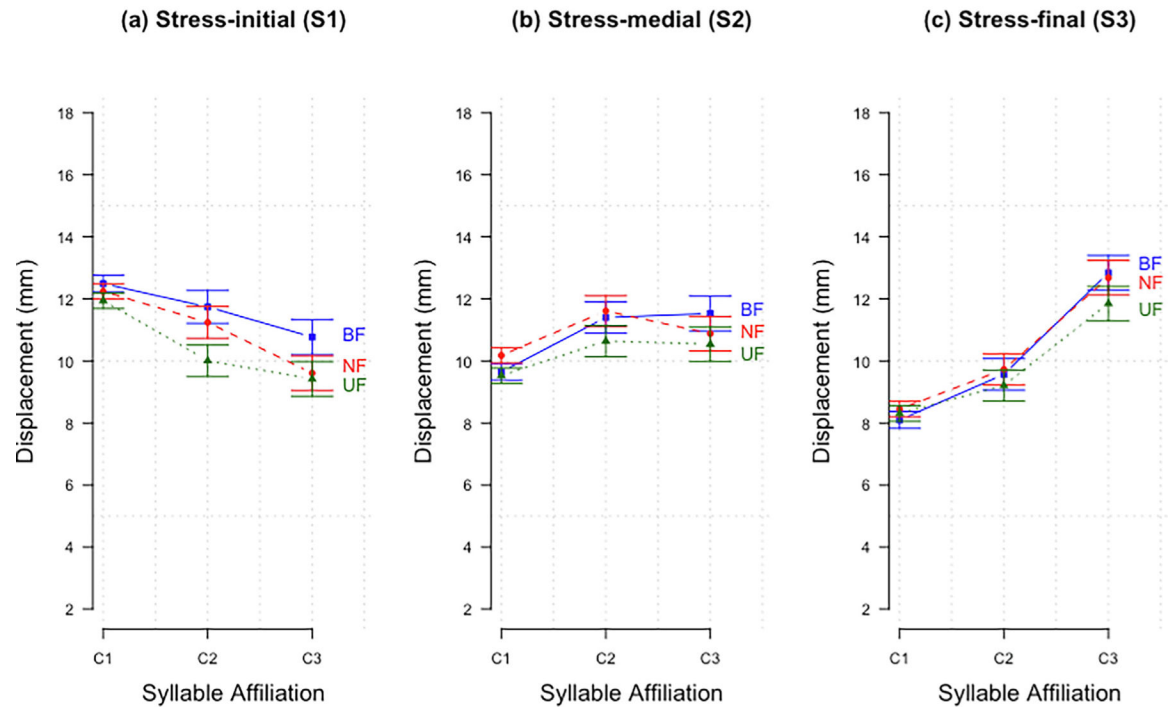


Fig. 17.

Displacement (in mm) of V gestures by syllable affiliation (V1, V2, V3) and focus type (BF, NF, UF) across peak velocity values per stress position (S1, S2, S3).

V Gestures

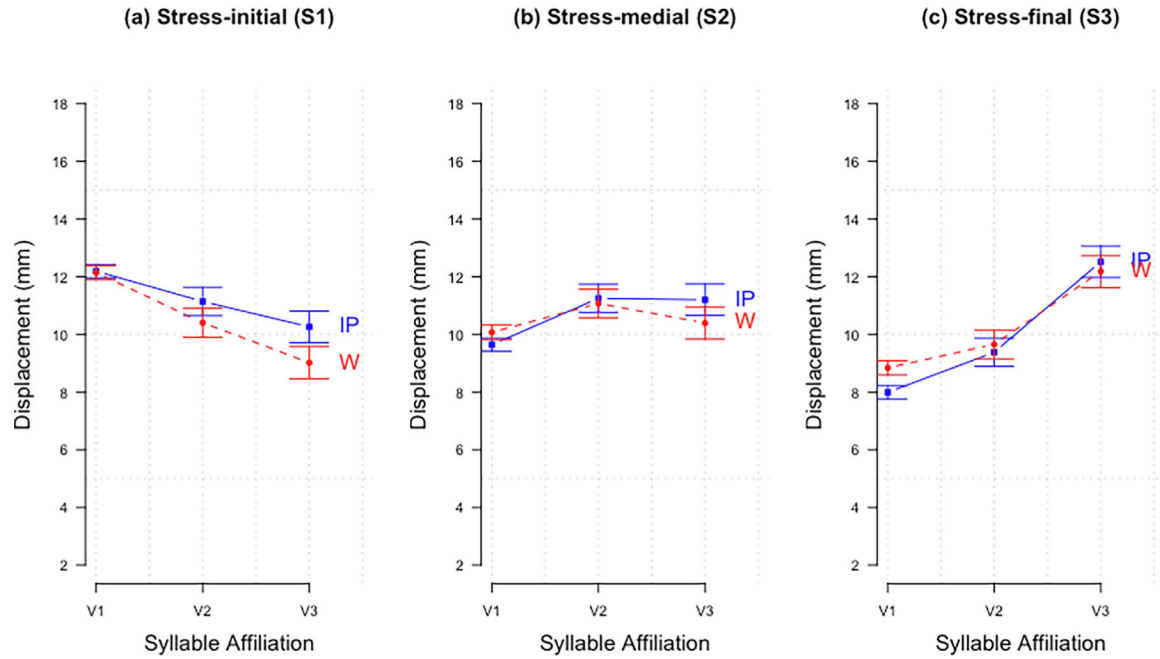


Fig. 18.

Displacement (in mm) of V gestures by syllable affiliation (V1, V2, V3) and boundary type (IP, W) across peak velocity values per stress position (S1, S2, S3).

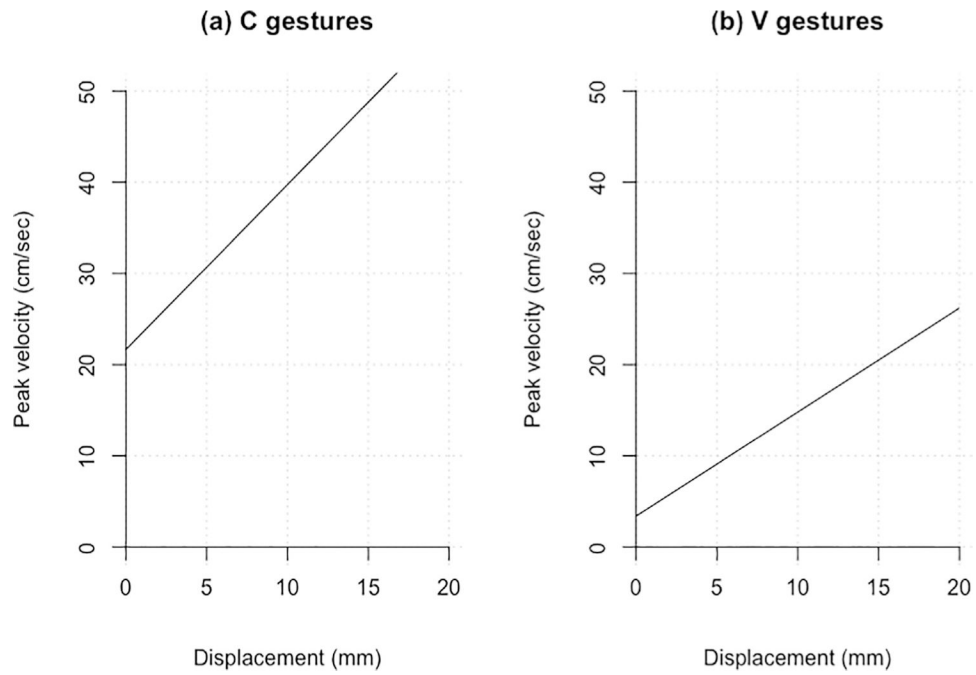


Fig. 19. Relationship between predicted peak velocity (in cm/sec) and displacement (in mm), based on the β estimate of displacement, for (a) C gestures and (b) V gestures.

C Gestures

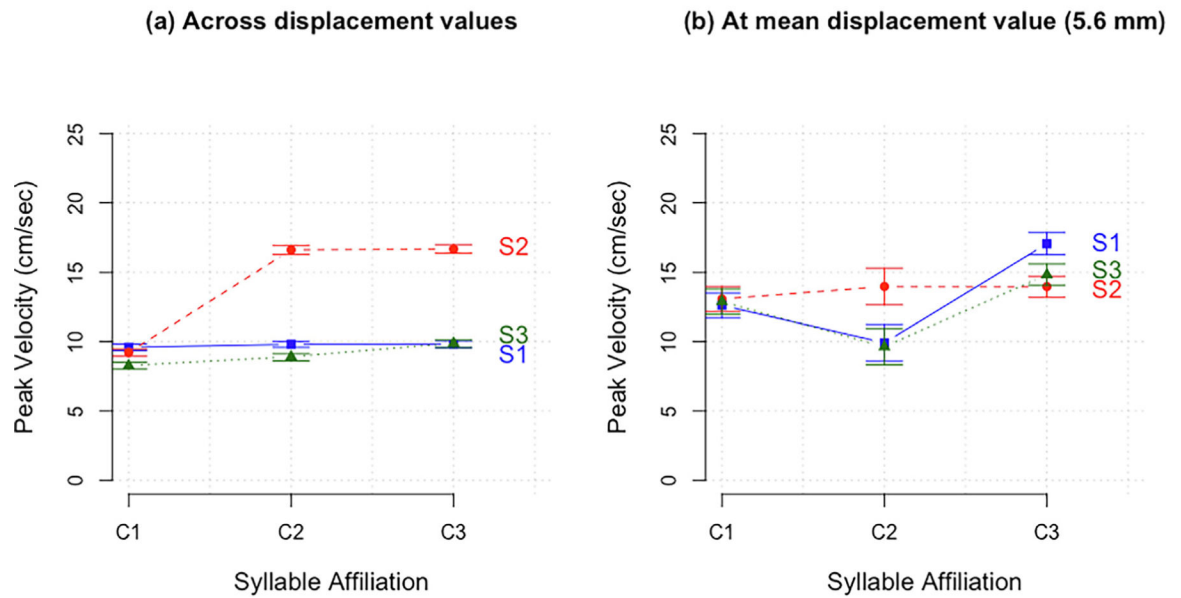


Fig. 20. Peak velocity (in cm/sec) of C gestures by syllable affiliation (C1, C2, C3) and stress position (S1, S2, S3) across displacement values in (a) and at mean value of displacement (5.6 mm) in (b).

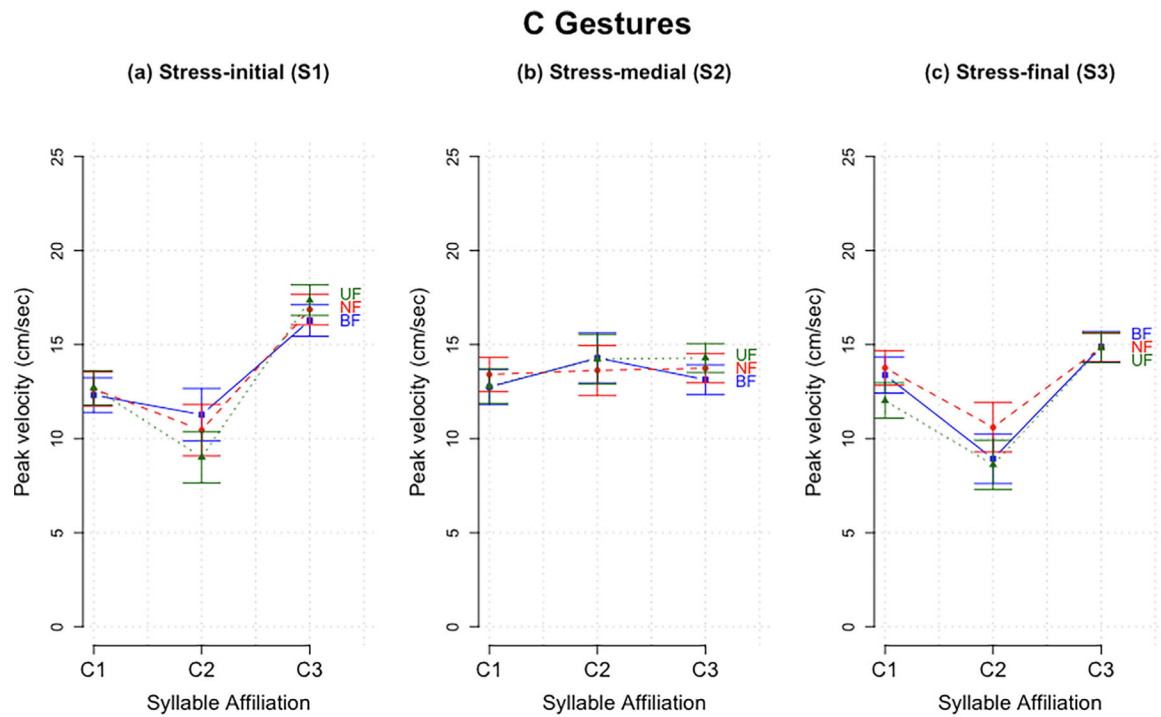


Fig. 21. Peak velocity (in cm/sec) of C gestures by syllable affiliation (C1, C2, C3) and focus type (BF, NF, UF) across displacement values per stress position (S1, S2, S3).

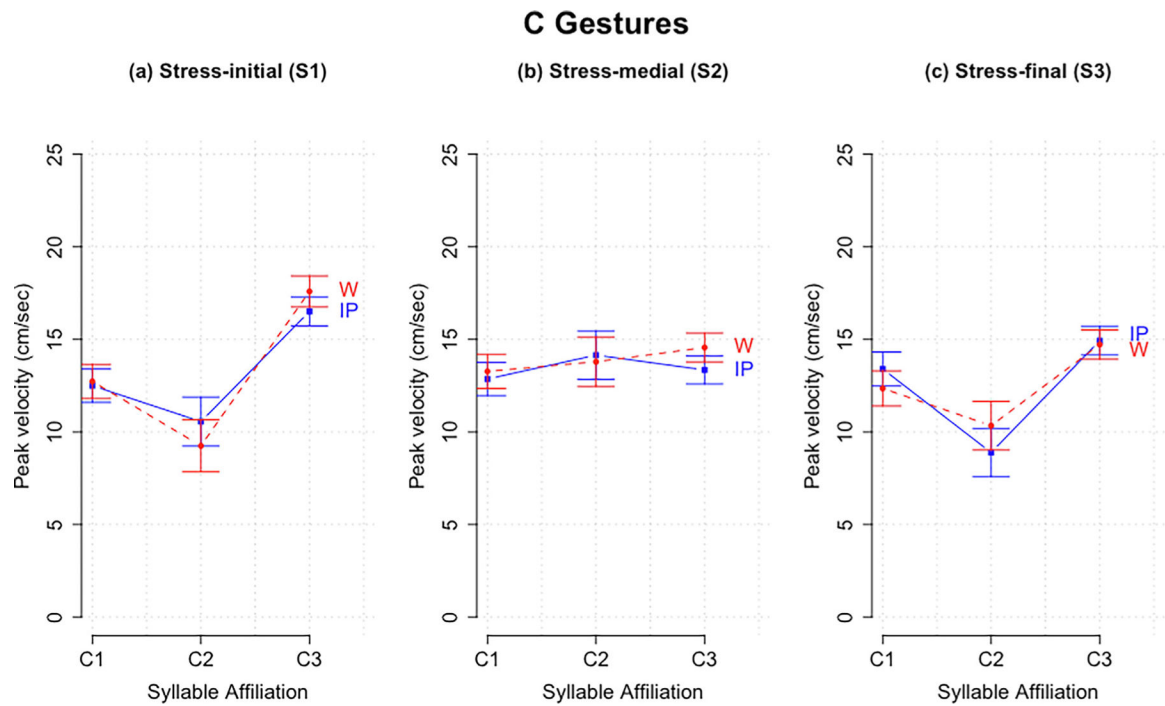


Fig. 22.

Peak velocity (in cm/sec) of C gestures by syllable affiliation (C1, C2, C3) and boundary type (IP, W) across peak velocity values per stress position (S1, S2, S3).

V Gestures

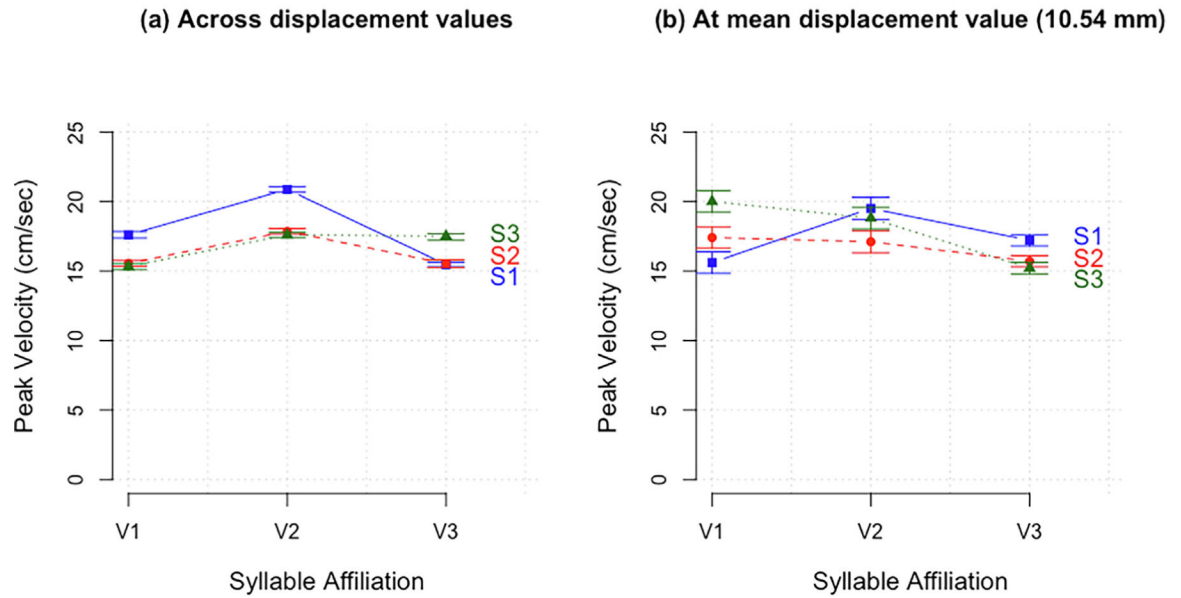


Fig. 23.

Peak velocity (in cm/sec) of V gestures by syllable affiliation (V1, V2, V3) and stress position (S1, S2, S3) across displacement values in (a) and at mean value of displacement (10.5 mm) in (b).

V Gestures

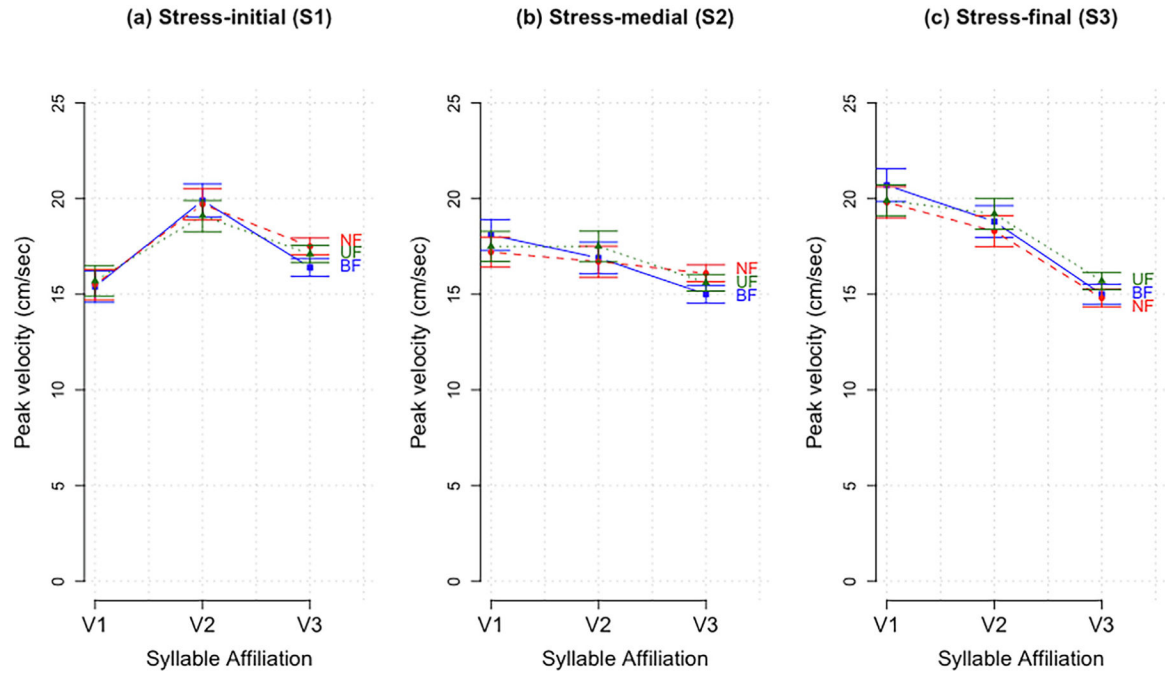


Fig. 24.

Peak velocity (in cm/sec) of V gestures by syllable affiliation (V1, V2, V3) and focus type (BF, NF, UF) across peak velocity values per stress position (S1, S2, S3).

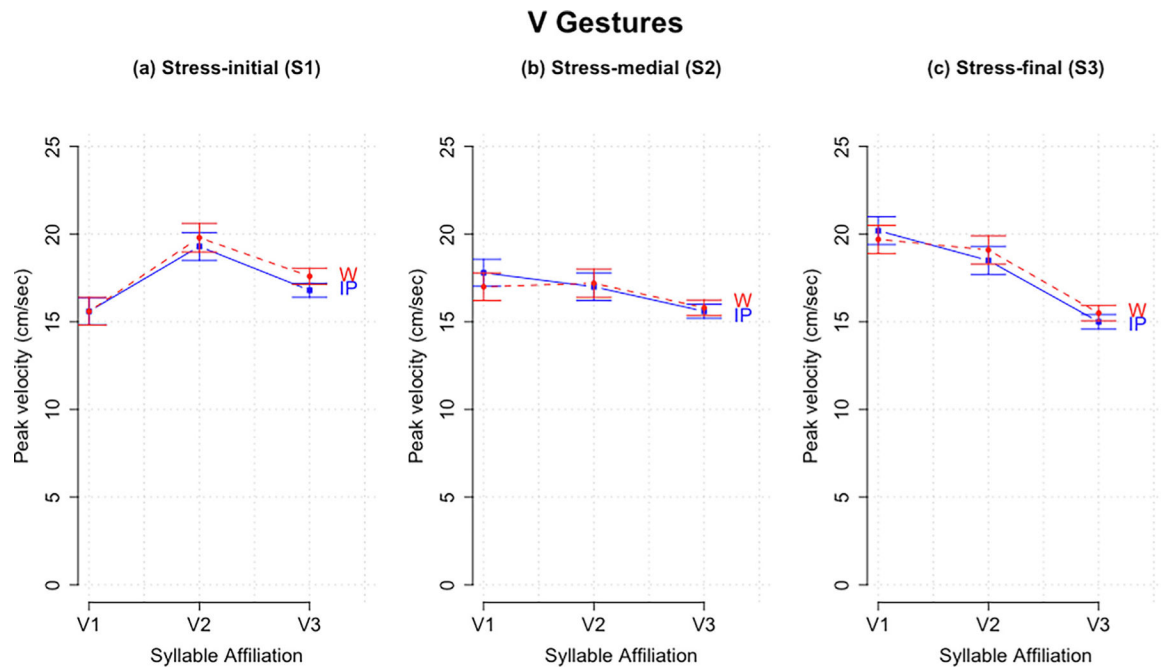


Fig. 25.

Peak velocity (in cm/sec) of V gestures by syllable affiliation (V1, V2, V3) and boundary type (IP, W) across peak velocity values per stress position (S1, S2, S3).

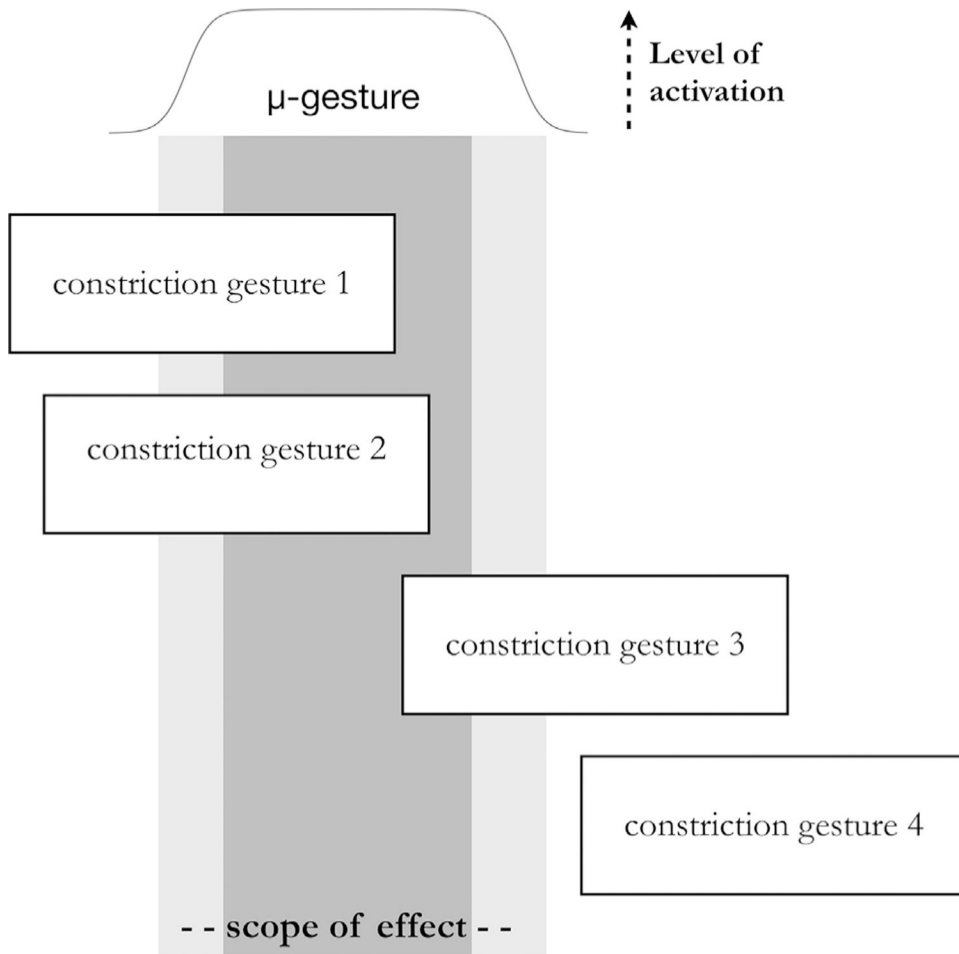


Fig. 26.

A schematic representation of a μ -gesture (adapted from Byrd and Saltzman (2003)). The two-shade gray box represents the scope of the modulation effects caused by the μ -gesture, with darker gray corresponding to the μ -gesture's maximal level of activation. Constriction gestures 1, 2 and 3, which are co-active with the μ -gesture, undergo the effects. Constriction gesture 4 does not overlap with the μ -gesture, and is thus unaffected.

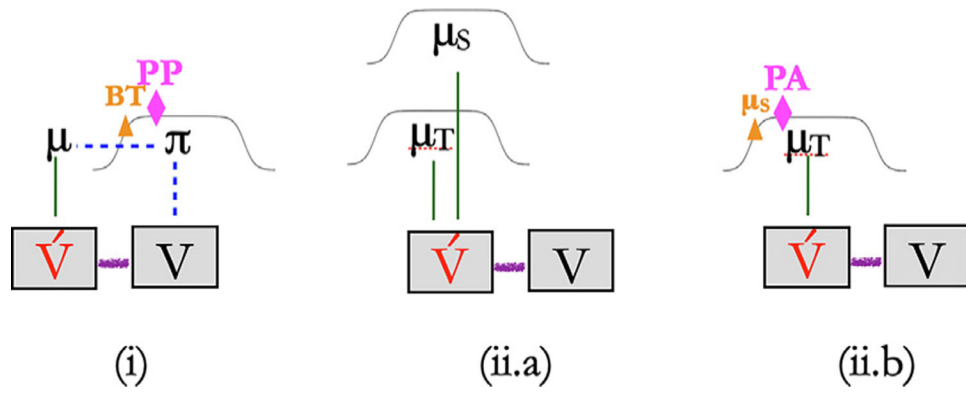


Fig. 27.

Schematic representation of a π - (i) and a μ -gesture (ii) and their coordination in the context of a stress-initial disyllabic word with two V gestures. Two alternative models for the μ -gesture coordination are shown (ii.a and ii.b). Solid lines represent in-phase coordination and broken lines represent anti-phase coordination. Stress is represented by ‘‘’. Triangles and white diamonds correspond to the strength levels of π - and μ -gesture activation that triggers the respective prosodic event. BT stands for boundary tone gesture (BT), PP for pause posture, PA for Pitch Accent gesture, μ_T for temporal μ -gesture and μ_S for spatial μ -gesture. This model is an extension of the proposal in Katsika (2016).

Table 1

Summary of Hypotheses.

HI: General kinematic profile of prominence	Longer, larger, and faster gestures in prominent vs. non-prominence positions
HII: Hierarchy of prominence	a. unstressed < stressed
	b. stressed < accented
	c. within accented: broad < narrow focus
HIII: Position of stress in the word	No effect
HIIV: Position of word in the phrase	Prominence * boundary interactions

Table 2

The stimuli for stress-initial words (MAMima). A rough translation into English of the context sentence (if present) is given first, and the target sentence in IPA along with a rough translation into English follows. The words bearing the nuclear pitch accent are marked with bold letters. Punctuation marks stand for phrase boundaries. Stimuli (1–4) correspond to the de-accented conditions, while stimuli (5–9) to the accented ones. In stimuli (1–3) and (5–8), the test word is phrase-final, while in stimuli (4) and (9) the test word is phrase-medial.

1	Negative Declarative Showing Reservation (ND): <i>What they are doing is horrible!</i> äe äjɛci'nun 'ɛkɔpi 'mɛmimɛ. mɛtɛ'ksi mɛθi'tɔn kɛrɛmɛ'litɛs pu'lun. It is not that they merchandize raw MAMima. It is just 'candies' they sell to students.
2	Wh-Question (WhQ): <i>We are looking for raw MAMima.</i> pu'psɛxnɛtɛ 'jɛkɔpi 'mɛmimɛ? mɛtɛ'ksi mɛθi'tɔn ɛ'vrɛɔs äjɛci'nite. Where are you looking for raw MAMima? Usually one can find some among students.
3	Imperative Request (IR): <i>You seem as if you want to ask me for a favor.</i> 'vrɛsmu 'liji 'ɛkɔpi 'mɛmimɛ. mɛtɛ'ksi mɛθi'tɔn ɛ'vrɛɔs äjɛci'nite. Find some raw MAMima for me. Usually one can find some among students.
4	Negative Declarative Showing Reservation, IP-medial (ND-medial): <i>What they are doing is unacceptable!</i> äe äjɛci'nun 'ɛkɔpi 'mɛmimɛ mɛtɛ'ksi mɛθi'tɔn kjɛ'nɪlikɔn ɛ'fivɔn. It is not that they merchandize raw MAMima to students and underage teenagers.
5	Affirmative Declarative (AD): <i>What are you looking for in our school?</i> ɛnɛzi'tɔ 'ɛkɔpi 'mɛmimɛ. mɛtɛ'ksi mɛθi'tɔn ɛ'vrɛɔs äjɛci'nite. I am looking for raw MAMima. Usually one can find some among students.
6	Yes-no question (YNQ): ɛnɛzi'tɛs 'ɛkɔpi 'mɛmimɛ? mɛtɛ'ksi mɛθi'tɔn ɛ'vrɛɔs äjɛci'nite. Are you looking for raw MAMima? Usually one can find some among students.
7	Causative clause (CC): ɛ'fu 'vrɪskun 'ɛkɔpi 'mɛmimɛ, mɛtɛ'ksi mɛθi'tɔn licɪu tin äjɛci'nun. Since it happens to have in their possession raw MAMima, they merchandize it to students.
8	Parenthetical Clause (PC): What do these people merchandize? fɪ'tɛ – 'mɛlɔn 'ɛkɔpi 'mɛmimɛ – mɛtɛ'ksi mɛθi'tɔn jimnɛ'siu äjɛci'nun. They merchandise plants – most likely raw MAMima – to high school students.
9	Affirmative declarative, IP-medial (AD-medial): What types of narcotic plants do they merchandize to students and underage teenagers? äjɛci'nunɛ 'ɛkɔpi 'mɛmimɛ mɛtɛ'ksi mɛθi'tɔn kjɛ'nɪlikɔn ɛ'fivɔn. They merchandize raw MAMima to students and underage teenagers.

Table 3

Significant fixed effects of model of C formation duration.

(1) Focus : Syllable affiliation	$F(4) = 3.48, p < 0.05$
(2) Focus : Stiffness	$F(2) = 4.97, p < 0.05$
(3) Boundary : Focus : Stress position	$F(2) = 2.4, p < 0.05$
(4) Boundary : Stress position : Syllable affiliation	$F(4) = 17.71, p < 0.05$
(5) Boundary : Stress position : Stiffness	$F(2) = 3.29, p < 0.05$
(6) Boundary : Syllable affiliation : Stiffness	$F(2) = 3.79, p < 0.05$
(7) Stress position : Syllable affiliation : Stiffness	$F(4) = 16.04, p < 0.05$

Table 4

Significant fixed effects of model of V formation duration.

(1) Focus : Stress position : Syllable affiliation	$F(8) = 2.49, p < 0.05$
(2) Focus : Syllable affiliation : Stiffness	$F(4) = 9.69, p < 0.05$
(3) Boundary : Focus : Stress position	$F(2) = 5.1, p < 0.05$
(4) Boundary : Stress position : Syllable affiliation	$F(4) = 5.84, p < 0.05$
(5) Boundary : Stress position : Stiffness	$F(2) = 6.62, p < 0.05$
(6) Stress position : Syllable affiliation : Stiffness	$F(4) = 8.34, p < 0.05$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 5

Summary of results on gestural duration.

Stiffness	Gestural duration decreased as stiffness (estimated as the ration of peak velocity/displacement) increased.
Stress	Stressed gestures were longer than unstressed gestures. Gestures of syllables adjacent to stressed ones were longer than those of syllables further away from the stress.
Focus and Stress	Regardless of stress position, V gestures of final syllable encoded focus type: BF > NF > UF. Other syllables, regardless of their stress status, distinguished focused from unfocused, and thus accented from de-accented.
Boundaries and stress	Boundary-related lengthening extended to the word-medial syllable in stress-initial and stress-medial words, and affected only the final syllable in stress-final words. In initial syllables, V gestures underwent boundary-related shortening regardless of their stress status, while C gestures presented shortening only in stress-initial words.

Table 6

Significant fixed effects of model of C gesture displacement.

(1) Stress position : Syllable affiliation: Peak Velocity	$F(8) = 130.5, p < 0.05$
(2) Focus : Syllable affiliation : Peak Velocity	$F(4) = 4.4, p < 0.05$
(3) Focus : Stress position : Syllable affiliation	$F(2) = 4.1, p < 0.05$
(4) Boundary : Syllable affiliation : Peak Velocity	$F(4) = 8, p < 0.05$
(5) Boundary : Stress position : Syllable affiliation	$F(2) = 9.7, p < 0.05$
(6) Boundary : Focus : Stress position	$F(4) = 3.3, p < 0.05$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 7

Significant fixed effects of model of V gesture displacement.

(1) Stress position : Syllable affiliation : Peak velocity	$F(8) = 161.6, p < 0.05$
(2) Focus : Syllable affiliation : Peak velocity	$F(4) = 62.3, p < 0.05$
(3) Focus : Stress position : Peak velocity	$F(2) = 110.9, p < 0.05$
(4) Focus : Stress position : Syllable affiliation	$F(4) = 48.5, p < 0.05$
(5) Boundary : Stress position : Peak velocity	$F(2) = 15.6, p < 0.05$
(6) Boundary : Focus : Peak velocity	$F(4) = 44.8, p < 0.05$
(7) Boundary : Syllable affiliation	$F(4) = 120.5, p < 0.05$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 8

Summary of results on gestural displacement.

Peak velocity	Gestural displacement increased as peak velocity increased.
Stress	Gestures of stressed syllables were larger than those of unstressed syllables. Spillover effects were strong, with C gestures showing largest displacement on the syllable following the stressed one. V gestures' largest displacement occurred in the stressed syllable. Weaker anticipatory effects were also detected.
Focus	Accent and/or focus increased the displacement of gestures in focused words, but did not make the stressed gestures even larger. Stress position and syllable affiliation affected the profile of these effects.
Boundaries	Final gestures of phrase-final words presented larger displacement and initial gestures smaller displacement than the respective gestures of phrase-medial words. The profile of boundary-related amplification was dependent on stress position in ways that further suggested significant spillover effects of stress.

Table 9

Significant fixed effects of model of C peak velocity.

(1) Stress position : Syllable affiliation : Displacement	$F(4) = 103.5, p < 0.05$
(2) Focus : Syllable affiliation : Displacement	$F(4) = 3.9, p < 0.05$
(3) Focus : Stress position : Displacement	$F(4) = 4.1, p < 0.05$
(4) Focus : Stress position : Syllable affiliation	$F(8) = 2.9, p < 0.05$
(5) Boundary : Stress position : Displacement	$F(2) = 7.9, p < 0.05$
(6) Boundary : Stress position : Syllable affiliation	$F(4) = 9.95, p < 0.05$
(7) Boundary : Focus	$F(1) = 5.5, p < 0.05$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 10

Significant fixed effects of model of V peak velocity.

(1) Stress position : Syllable affiliation : Displacement	$F(4) = 14.9, p < 0.05$
(2) Focus : Syllable affiliation : Displacement	$F(4) = 19.4, p < 0.05$
(3) Focus : Stress position : Displacement	$F(4) = 8.3, p < 0.05$
(4) Focus : Stress position : Syllable affiliation	$F(8) = 2.4, p < 0.05$
(5) Boundary : Syllable affiliation : Displacement	$F(2) = 10.1, p < 0.05$
(6) Boundary : Stress position : Displacement	$F(2) = 20.9, p < 0.05$
(7) Boundary : Focus : Displacement	$F(1) = 18.3, p < 0.05$

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 11

Summary of results on gestural displacement.

Displacement	Gestural peak velocity increased as displacement increased.
Stress	There were no systematic effects of stress on gestural velocity independently of displacement.
Focus	There were no systematic effects of focus/accent on gestural velocity independently of displacement.
Boundaries	There were no systematic interaction effects between prominence (stress and/or accent) and boundary type on gestural velocity independently of displacement.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript