

Tradición y Diversidad en los Aspectos Psicológicos, Socioculturales y Musicológicos de la Formación Musical. Conservatorio Municipal de Bahía Blanca, Bahía Blanca, 2010.

Validación de restricciones sobre un modelo generativo para el establecimiento de implicaciones melódicas.

Pablo Rodríguez Zivic y Favio Shifres.

Cita:

Pablo Rodríguez Zivic y Favio Shifres (Mayo, 2010). *Validación de restricciones sobre un modelo generativo para el establecimiento de implicaciones melódicas. Tradición y Diversidad en los Aspectos Psicológicos, Socioculturales y Musicológicos de la Formación Musical. Conservatorio Municipal de Bahía Blanca, Bahía Blanca.*

Dirección estable: <https://www.aacademica.org/favio.shifres/139>

ARK: <https://n2t.net/ark:/13683/puga/sOm>

Acta Académica es un proyecto académico sin fines de lucro enmarcado en la iniciativa de acceso abierto. Acta Académica fue creado para facilitar a investigadores de todo el mundo el compartir su producción académica. Para crear un perfil gratuitamente o acceder a otros trabajos visite: <https://www.aacademica.org>.

VALIDACIÓN DE RESTRICCIONES SOBRE UN MODELO GENERATIVO PARA EL ESTABLECIMIENTO DE IMPLICACIONES MELÓDICAS

PABLO RODRÍGUEZ ZIVIC* Y FAVIO SHIFRES**

*UNIVERSIDAD DE BUENOS AIRES

**UNIVERSIDAD NACIONAL DE LA PLATA

Fundamentos

En los últimos 30 años ha habido una notable contribución a la psicología de la música desde teorías musicológicas que se encuentran a un paso de ser computacionales. Ejemplos de esto son la *Teoría Generativa de la Música Tonal* de Lerdahl y Jackendoff (1983), la teoría de la *Implicación-Realización* de Narmour (1990, 1992) y la teoría del *Espacio Tonal de Alturas* de Lerdahl (2001). Asimismo, otras teorías psicológicas se apoyan fuertemente en principios de la tradición teórico occidental, como la teoría de la estabilidad tonal de Krumhansl (1990), y la teoría del anclaje melódico de Bharucha (1984). Se propone entonces utilizar estas teorías entre otras para construir un modelo generativo/computacional para líneas melódicas con particular cuidado en la aplicación de los principios que ellas proponen.

Este modelo se encuentra conformado por pequeñas unidades responsables de la generación de ciertos atributos de la melodía resultante. Cada unidad se encuentra conformada por un modelo probabilístico y un mecanismo de interacción con el resto de las unidades. Este mecanismo permite dos cosas: por un lado *entrenar* el modelo probabilístico a partir de una partitura, y por el otro utilizarlo para *generar* instancias del atributo musical que este modela. El hecho de que el mecanismo de interacción se abstraiga del modelo que utiliza cada unidad generativa permite aislar cada unidad del resto y controlar su comportamiento de forma individual.

Utilizando un modelo de las características mencionadas es posible construir composiciones de forma automática en donde se varíe deliberadamente el comportamiento de una cierta unidad sin interferir el comportamiento del resto para luego analizar el impacto de tal variación en las composiciones resultantes.

Una de las unidades generativas que el modelo implementa se encuentra basada en dos pilares teóricos: 1) por un lado, la teoría de la Implicación-Realización propuesta por Leonard B. Meyer (1956, 1973) y sistematizada por Eugene Narmour (1977; 1990; 1992), brinda una distribución de probabilidad que permite al programa elegir una nota de la línea melódica teniendo en cuenta la dirección y el tamaño del intervalo anterior; 2) por el otro, la inducción de la tonalidad en la que el programa tiene que componer la melodía da lugar a una distribución de probabilidad que tiene en cuenta la distribución estadística de los eventos tonales del input (la *event hierarchy*, Bharucha y Krumhansl, 1983) obtenida y exhibida como el *perfil de alturas (pitch profile)* de acuerdo al algoritmo propuesto por Krumhansl (1990). De allí, esta unidad implementa una distribución de probabilidad, que dado el último intervalo melódico generado, establece cuál es la probabilidad del siguiente intervalo. Obsérvese que en este caso, al estar el primer intervalo melódico ya establecido, elegir el segundo intervalo corresponde a elegir solamente una nota, puesto que la primera nota del intervalo a ser elegido está dada. Así, ambos componentes – el melódico interválico, modelizado por el la teoría de implicación-realización de Narmour, y el melódico tonal, en base a la inferencia de la “Jerarquía de Eventos” (Bharucha y Krumhansl 1983) extraída en forma de pitch profiles (Krumhansl 1990) – conforman una única distribución de probabilidad.

Es importante notar que existe una interpretación dual para estos dos modelos que operan en conjunto. Ambos modelos son probabilísticos, y de esta forma, una primera interpretación es esa: los eventos que tengan asociado un valor más alto serán más probables que los que tengan asociados un valor menor. Sin embargo, dado que el modelo original propuesto por Narmour predica sobre el grado de expectativa que generan ciertos intervalos melódicos sobre las posibles continuaciones, el modelo construido puede interpretarse de forma tal que las continuaciones más probables dado el contexto también sean las más esperadas. Respecto al modelo de configuración tonal, la interpretación estará atada al grado de estabilidad de una nota: una nota que reciba un valor

mayor será más estable. De esta forma, en el modelo compuesto, la interpretación será que las notas más probables son fuertemente implicadas y además estables.

Por sobre esta unidad generativa se construyó otra con el objetivo de configurar morfológicamente la melodía como una cadena de unidades discursivas. Para esto se consideró la noción de frase descrita por Westergaard (mencionada en Rothstein 1989). Esta aproximación teórica establece que la frase musical consta de dos conjuntos de alturas, que se vinculan como tal en función de tres propiedades:

1. Dado el primer conjunto de alturas, uno espera el segundo.
2. Uno tiene alguna noción de cuando este conjunto ocurrirá
3. Una vez que el segundo conjunto ocurrió, uno constata que la frase llegó a su meta

En analogía, se elaboró una definición operativa de unidad discursiva con el objetivo de computar configuraciones morfológicas (unidades de discurso) que respeten tales criterios. De acuerdo con esta definición operativa dos acordes sucesivos son tomados como los “dos conjuntos de alturas”. Por lo tanto la noción de unidad formal está anclada al ritmo de la función armónica. De este modo la unidad generativa respeta la primera restricción para la configuración morfológica propuesta por Westergaard. Las otras dos restricciones se aplicaron valiéndose del modelo compuesto para la elección de las notas melódicas arriba. La noción de meta está dada el grado de estabilidad de la nota a la que se arriba. Esta nota se denominó *nota de apoyo*, y la previsión del momento de arribo a dicha nota derivaría de las restricciones que dicha nota impone sobre las dos notas anteriores de modo de que el arribo a dicha meta sea fuertemente implicado.

La fortaleza de esta implicación requiere que se establezcan restricciones sobre la probabilidad asignada originalmente por el modelo. Para ello se decidió utilizar percentiles que restrinjan la distribución de probabilidad que se utilizará. De esta forma si el percentil es un P%, sólo se utilizará el P% más probable dentro de la distribución. Así, si el percentil es 1 se utilizará el 100% de la distribución, si el percentil es 0,5 se utilizará el 50% más probable de la distribución, y así sucesivamente. Aquí, para el caso en el caso de que P tenga el valor 0, se optó por cambiar el significado, y en vez de dejar la distribución vacía, se utilizará solamente el evento más probable.

Dado que en la definición operativa propuesta para la generación de unidades discursivas utiliza como noción de meta la nota de apoyo del acorde siguiente, lo primero que debe hacerse antes de componer una frase es elegir la nota de apoyo a la que se implicará.

En la Figura 1 se exhiben los compases 6, 7 y 8 de la melodía generada a partir de la Danza Alemana WoO 13 N 11 de Beethoven utilizando el percentil 0. Supóngase que se ha generado la melodía hasta donde se muestran los signos de interrogación. La nota que se encuentra luego de los signos de interrogación es la nota tomada como punto de apoyo para la frase siguiente.



Figura 1. Un fragmento de una melodía donde ya se eligió la próxima nota (Mi) de apoyo y restan tocar dos notas.

La meta de que la nota Mi sea fuertemente implicada hace que se establezcan restricciones sobre las dos notas que faltan tocar. Siendo que estas restricciones están basadas en la interpretación dual del modelo compuesto, en la figura 2 se presenta cual sería la probabilidad de la nota de apoyo (Mi) para las distintas posibilidades de las notas que están marcadas con signo de interrogación. Como el percentil utilizado en este caso es el percentil 0, sólo las notas de la zona marcada con rojo en la figura 2 pueden ser utilizadas antes que la nota de apoyo. Asimismo, estas notas, al proceder de un cierto contexto melódico (representado por las dos notas anteriores, La y Mi), deberán ser elegidas también a partir de la distribución de probabilidad que el modelo compuesto instancie en tal contexto (las notas La y Mi).

La distribución de probabilidad presentada en la figura 3, no corresponde al modelo compuesto solamente, sino que por sobre este se le aplican las restricciones exhibidas en la figura 2, de esta forma sólo las notas que pertenecen a la zona marcada con rojo en la figura 2 tienen probabilidad distinta de cero en la figura 3.

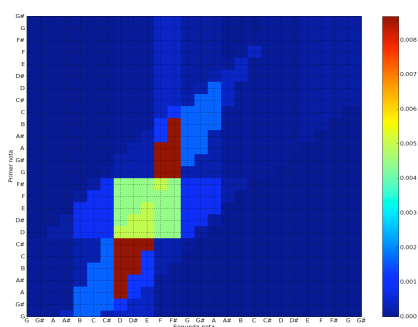


Figura 2. Probabilidad de la nota de apoyo (Mi) para las dos notas anteriores. En el eje X se muestra la primer nota, y en el eje Y se muestra la segunda

Nótese que si bien se podría considerar que los valores numéricos de las probabilidades son bajos, hay que tener en cuenta que lo importante es la proporción que hay entre estos valores. Es decir, en el gráfico se puede ver que las dos continuaciones más probables son o bien tocar las notas Do# y luego Re o bien tocar las notas La# y luego Fa#. Estas dos continuaciones tienen cada una una probabilidad aproximadamente de 1.37×10^{-5} y 4.51×10^{-6} , que si bien en términos absolutos es un valor pequeño, la continuación más probable es aproximadamente 3.04 más probable que la que le sigue.

La decisión que la construcción del programa demanda en relación al percentil aplicado para esta restricción es la de determinar cuál es el percentil adecuado para permitir la variedad melódica sin perder la estabilidad tonal. Para ello se consideró pertinente relevar evidencia psicológica a través de juicios de “buena conformación” brindados por oyentes enculturados, que permita validar la elección de un determinado percentil.

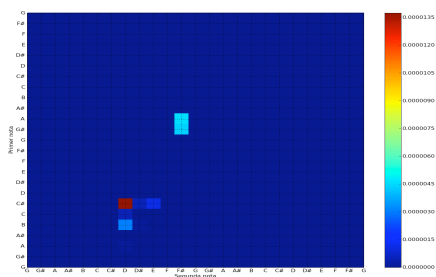


Figura 3. Distribución de probabilidades que explica la elección de las dos notas. El eje X representa la primer nota, el eje Y representa la segunda, y el color representa la probabilidad en la escala de colores de la figura

Objetivos

Este trabajo se propone testear la plausibilidad de las melodías generadas de acuerdo a diferentes percentiles con el objeto de encontrar un valor de percentil adecuado para restringir la distribución de probabilidad sobre la que el programa selecciona el intervalo implicativo a las notas de apoyo.

Método

Sujetos

43 personas (20 mujeres y 23 varones) fueron convocadas voluntariamente por correo electrónico para realizar la tarea. La media de edad fue de 34,23 años (mínima 19 y máximo 64 años). Fueron agrupadas de acuerdo con tres niveles de formación musical sistemática: No músicos

(sin formación musical; 16 sujetos), Formación Moderada, (hasta 10 años de formación musical sistemática; 11 sujetos), Músicos profesionales (más de 10 años de formación musical sistemática; 16 sujetos).

Estímulos

Se seleccionaron 5 fragmentos de piezas para piano del repertorio académico clásico: (1) Danza Alemana WoO 13 N 11 de Beethoven; (2) Contradanza WoO 14 N 2 de Beethoven; (3) Contradanza WoO 14 N 8 de Beethoven; (4) Danza Alemana D973 N 1 de Schubert; y (5) Danza Alemana D974 N 2 de Schubert. Estas piezas fueron utilizadas como input del programa. Cada uno de ellos generó 4 melodías de la extensión de la pieza original (aproximadamente 30 segundos) de acuerdo a 4 valores diferentes para el percentil que restringe la probabilidad de que las dos notas anteriores a cada punto de apoyo sean elegidas de acuerdo al modelo de Implicación-realización melódica de E. Narmour. Los valores del percentil elegidos fueron 0; 0,3; 0,8 y 1. De tal modo en las melodías generadas con la indicación del percentil 0 las notas de apoyo eran más probablemente abordada desde un par de notas anteriores siguiendo la teoría de Narmour. Por el contrario, las melodías generadas con el percentil 1 no presentaban para nada esa restricción. Los restantes componentes del modelo de generación de melodías fueron mantenidos fijos para todas las melodías. Las melodías fueron ejecutadas por la computadora con valores de duración y velocidad (sonoridad) normalizados de acuerdo a lo estipulado en la partitura (no expresivo), con un timbre de piano.

Aparatos

Los estímulos fueron desarrollados utilizando *La compositora* (Rodríguez Zivic 2009). La prueba fue autosuministrada a través de una página web, de tal modo que cada sujeto utilizaba sus propios recursos de reproducción de sonido.

Procedimiento

Los sujetos escuchaban cada una de las 20 melodías y tenían que evaluarlas de acuerdo a sus propios criterios de “buena forma” en una escala de 5 puntos. Los sujetos entraban a la página web a través de un link que les era suministrado vía correo electrónico. Allí se les comunicaba los fundamentos y objetivos de la prueba y se les proporcionaba ejemplos de la tarea de composición del programa. A continuación se les explicaba el procedimiento y se le proporcionaban 4 ejemplos de aprestamiento. Seguidamente se sucedían los 20 ítems de la prueba. Cada sujeto determinaba el momento de escuchar cada ejemplo clickeando sobre un ícono de “play”, y colocaba su puntuación directamente en un dispositivo ad hoc diseñado en la página. Al finalizar la prueba respondían un cuestionario de datos personales.

Diseño

Las 20 melodías fueron suministradas en orden aleatorio para cada sujeto.

Resultados

Se ejecutó un modelo lineal general de mediciones repetidas tomando la pieza input (5 niveles) y los percentiles de acuerdo a los que se generaron las melodías target (4 niveles) como factores intra-sujetos. Se consideraron los grupos de experiencia musical como factor entre-sujetos.

Solamente el factor *PIEZA* resultó significativo ($F_{[4-39]}=4.429$; $p=.002$). Lo que más interesa aquí, el factor *PERCENTIL* arrojó una significación marginal ($F_{[3-40]}=2.585$; $p=.056$), sin embargo un contraste *post hoc* mostró que el percentil 1 fue evaluado significativamente diferente del percentil 0 ($F_{[1-42]}=5.630$; $p=.023$) y del percentil 0,3 ($F_{[1-42]}=4.029$; $p=.052$). El gráfico de la figura 4 muestra esas diferencias.

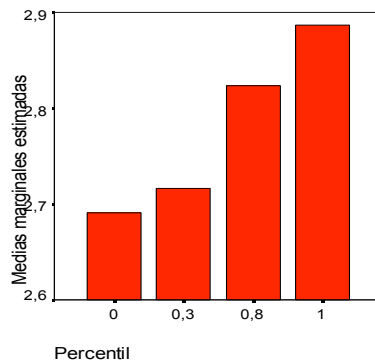


Figura 4. Medias de los juicios de los sujetos para los 4 percentiles propuestos para la restricción

Paradójicamente, los sujetos consideraron mejores las melodías que no estaban restringidas (o que solo lo estaban moderadamente) por el modelo compuesto (Narmour-Krumhansl).

La interacción entre *PIEZA* y *PERCENTIL* también resultó significativa ($F_{[12-31]}=2.937$; $p=.001$). El gráfico de la figura 5 permite observar que solamente la primera pieza cumple moderadamente con la predicción que avala la teoría de la implicación interválica.

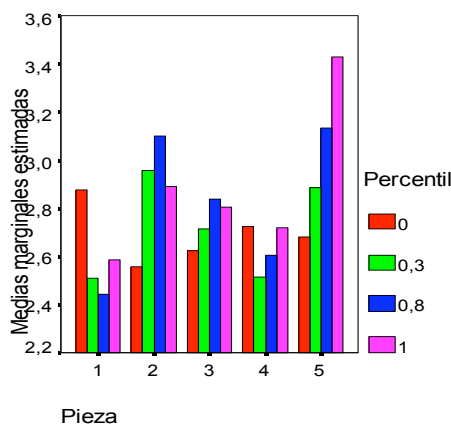


Figura 5. Interacción *PIEZA* – *PERCENTIL*.

Finalmente la interacción *PIEZA* y *EXPERIENCIA MUSICAL* arrojó una significación marginal ($F_{[8-35]}=2,107$; $p=.038$).

Discusión

El experimento reportado en este trabajo buscaba obtener evidencia empírica acerca de la importancia de que la unidad relativa a la selección de alturas para la melodía de un modelo generativo/computacional estuviese más o menos restringido por las teorías de implicación realización –que predice qué intervalos serán los más esperados– y de estabilidad tonal – que predice qué notas serán consideradas como más estables. La predicción decía que cuanto más alta fuera la restricción, los juicios de buena forma de los oyentes serían más altos. De este modo la restricción estipulada en el percentil 0 daría lugar a juicios más altos que la correspondiente al percentil 1. Sin embargo, los resultados parecen contradecir esta predicción. En esta sección se discuten algunos puntos que pueden echar luz sobre esta contradicción.

En lo metodológico

El experimento buscaba comparar la bondad de conformación de las melodías generadas, sin embargo la tarea experimental no era de comparación. Por el contrario la tarea era de juicio de “buena forma” para cada melodía en forma independiente. De tal manera, la perspectiva psicológica de la comparación no está capturada en esta metodología. Es posible pensar, entonces, que una

tarea de comparación en relación a la “buena forma” podría orientar al sujeto en la audición de los componentes que en efecto varían de una melodía a la otra -en este caso restringidos por la variable independiente (percentil).

En lo musicológico

Al observar en detalle la interacción entre los factores *PIEZA* y *PERCENTIL* se puede apreciar que la diferencia en los juicios relativos a las melodías generadas de acuerdo a los diferentes niveles de restricción varía de acuerdo a la pieza input. Por ejemplo las melodías generadas a partir de la pieza 1 dieron lugar a juicios cuya distribución es contraria a la correspondiente a los juicios para las melodías generadas por la pieza 5. Un análisis musical de las piezas compuestas por el programa puede entonces brindar algún indicio de las causas de estos resultados.

La figura 6 muestra la partitura de la pieza input 1 y las de las melodías generadas conforme los percentiles 0 y 1. En lo concerniente a las relaciones interválicas, se observa que el contorno melódico de la melodía percentil 0 es más suavizado (con menos y menores saltos), posibilitando una mayor previsión de las metas melódicas. En lo relativo a la conformación de la estabilidad tonal, se observa que ambas melodías se configuran tonalmente alrededor de si menor – a pesar de que la pieza está en La Mayor -. Mientras que la melodía percentil 0 se orienta hacia La Mayor a partir del compás 11, la melodía percentil 1, permanece en si menor hasta el final, de tal suerte que el La final no puede ser interpretado como tónica. Por el contrario el Do final de la melodía percentil 0 puede ser interpretado como la tercera de la función tónica con la correspondiente estabilidad como para garantizar un cierre y por ende mejorar la “buena conformación”: Es probable que ambos rasgos – el melódico y el tonal - estén favoreciendo las mayores puntuaciones de los oyentes para dicha melodía.

The image displays a musical score for 'Danza Alemana WoO 13 N 11' by Beethoven. It consists of six systems of music. The first system is a piano accompaniment, with a treble clef on the upper staff and a bass clef on the lower staff, labeled 'Piano'. The second system is another piano accompaniment, starting at measure 9. The remaining four systems are single melodic lines in a treble clef, representing generated melodies for percentiles 0 and 1. The key signature is three sharps (F#, C#, G#) and the time signature is 3/4.

Figura 6. *Pieza input N° 1: Danza Alemana WoO 13 N 11 de Beethoven (arriba) y las melodías generadas de acuerdo a las restricciones correspondientes al percentil 0 y al percentil 1 (de arriba hacia abajo)*

La figura 7 muestra la partitura original de la pieza input N°5 con las melodías generadas correspondientes a los percentiles 0 y 1 respectivamente. En lo referente a lo interválico, se aprecia en la melodía restringida al percentil 0, un contorno melódico más suavizado que el de la otra melodía particularmente en los primeros 8 compases y sobre la cadencia final, conforme a lo predicho por la teoría de Narmour (en relación al tamaño y la direccionalidad de los intervalos). En lo relativo a la configuración de la tonalidad ambas melodías están claramente planteadas en Reb Mayor (la tonalidad de la pieza input). Sin embargo la primera termina sobre la dominante apareciendo como inconclusa (inestable), mientras que la segunda termina sobre la tónica, precedida por la sensible, de modo que aparece como más estable. Parece ser, entonces, que los oyentes privilegian la estabilidad tonal a la implicación melódica a la hora de establecer sus juicios de “buena conformación” (varios sujetos dejaron consignadas observaciones de esta índole al finalizar el test).

La naturaleza compuesta del modelo de esta unidad generativa hace que no sea posible separar claramente un componente del otro para testarlos por separado. Sin embargo, a la luz de estos resultados aparece como importante obtener esos datos que permitan establecer la bondad de cada componente individualmente. En tal sentido es posible que eliminando el componente tonal, al seleccionar input atonales, se pueda evaluar mejor la modelización del componente interválico.

Piano

The image displays a musical score for a piano piece. It consists of six staves. The top two staves are a grand staff (treble and bass clefs) labeled 'Piano', showing the original melody and accompaniment. The bottom four staves are single treble clef staves, each starting with a '9' in the margin, representing generated melodies for different percentiles. The key signature is three flats (B-flat, E-flat, A-flat) and the time signature is 3/4. The original melody is in a 3/4 time signature, and the generated melodies are also in 3/4 time.

Figura 7. Pieza input N° 5: Danza Alemana D 974 N° 2 de Schubert (arriba) y las melodías generadas de acuerdo a las restricciones correspondientes al percentil 0 y al percentil 1 (de arriba hacia abajo)

En lo computacional

Como se vio, del análisis musicológico se desprende que es posible que los dos componentes del modelo requieran un tratamiento separado. Esto implicaría diseñar unidades independientes para ellos.

Pero además se observa a raíz de la desconfiguración tonal de las melodías correspondientes al input 1, que es posible que el modelo utilizado para generar la estabilidad tonal (Krumhansl) deba “componerse” con el de la frase melódica para determinar más ajustadamente las notas de apoyo de modo de garantizar una inducción tonal más ajustada.

Es posible también que la falla en la predicción se deba también a la velocidad del ritmo armónico. Nótese que en el fragmento donde la predicción (que el percentil 0 iba a recibir un puntaje mayor) fue acertada, figura 6, tiene un ritmo armónico que permite que se configuren unidades discursivas de varias notas. Sin embargo, en el fragmento correspondiente a la figura 7, el ritmo armónico es más “acelerado” en este sentido. Es posible que la solución a esto sea elaborar unidades discursivas que no necesariamente utilicen acordes sucesivos para los puntos de apoyo.

En todo caso, se considera necesario avanzar en la validación de las decisiones teóricas tomadas para la construcción del programa, sobre la base de la evidencia empírica obtenida de los juicios de los oyentes como se ha hecho en esta oportunidad.

Referencias

- Bharucha, J. J. (1984). Anchoring effects in music: The resolution of dissonance. *Cognitive Psychology*, **16**, 485-518.
- Bharucha, J. J. y Krumhansl, C. L. (1983). The representation of Harmonic structures in music: Hierarchies of stability as a function of context. *Cognition*, **13**, 63-102.
- Krumhansl, C. L. (1990). *Cognitive Foundations of Musical Pitch*. Oxford: University Press.
- Lerdahl, F. y Jackendoff, R. (1983). *A Generative Theory of Tonal Music*. Cambridge, MA: The MIT Press.
- Meyer, L. B. (1956). *Emotion and Meaning in Music* Chicago: The University Press.
- Meyer, L. B. (1973). *Explaining Music*. Chicago: The University Press.
- Narmour, E. (1977). *Beyond Schenkerianism. The need for Alternatives in Music Analysis*. Chicago and London: University of Chicago Press.
- Narmour, E. (1990). *The analysis and cognition of basic melodic structures: The implication-realization model*. Chicago: University of Chicago Press.
- Narmour, E. (1992). *The analysis and cognition of melodic complexity*, Chicago, University of Chicago Press.
- Rothstein, W. (1989). *Phrase Rhythm in Tonal Music*. New York y Londres. Schirmer Books.